SPATIAL POPULATION PROCESSES

by

ERIC RENSHAW

Thesis Submitted for the Degree of Doctor of Philosophy

University of Edinburgh

March 1976

DECLARATION

The following record of research work is submitted as a Thesis for the Degree of Doctor of Philosophy of the University of Edinburgh, having been submitted for no other Degree.

The research work was carried out under the supervision of Mr P.R. Fisk. Except where due acknowledgement is made, the work is original.

Eric Renshaw

Department of Statistics,

University of Edinburgh.

March 1976.

ABSTRACT OF THESIS

This thesis is a theoretical study of the effect of migration between colonies, each of which is developing according to a simple stochastic birth-death-immigration process.

In Chapters 2 to 7 I investigate the probability structure of the two-colony process. The Kolmogorov forward differential equation for the population size probabilities is developed and from it expressions are derived for the first- and second-order moments. Exact solutions to this forward equation are obtained for three special cases and a recursive solution is developed in a fourth. Three approximate solutions are developed; (i) by modifying the birth mechanism, (ii) by fitting a bivariate negative binomial distribution, and (iii) by placing an upper bound on the total population size. Iterative solutions are then derived by the use of two different techniques. In the first a power series solution is obtained in terms of a common migration rate. In the second sequences of functions are generated which converge to the required solution. The investigation of the two-colony process concludes with a simulation study and an analysis of the probability of extinction.

In Chapter 8 I introduce a 'stepping-stone' model in which the population is composed of an infinite number of colonies which may be considered to be situated at the integer points of a single co-ordinate axis. Migration is allowed between nearest-neighbours only. Although the Kolmogorov forward differential equation cannot

be solved directly, approximate solutions are developed in an analogous manner to those derived for the two-colony process. First- and second-order moments are obtained and an exact stochastic solution is developed for one special case.

If the population has a positive rate of growth and is initially concentrated into a relatively small geographic region, we may expect it to diffuse into the surrounding areas and eventually to take over the entire territory. This expanding population may be envisaged as generating a travelling wave and in Chapter 9 I investigate the velocity of propagation and the form of the wave profile.

In Chapter 10 I examine non-nearest-neighbour migration models and develop expressions for the mean size of each colony at time t for several appropriate migration distributions. To conclude the thesis I present a spatial model in two-dimensions and relate it to data on the spatial distribution of flour beetles in a closed container.

ACKNOWLEDGEMENTS

I should like to record my gratitude to Mr P.R. Fisk for his guidance, encouragement and helpful criticisms throughout this period of research.

My thanks are also due to Dr D. Mollison for helpful discussions relating to the spatial spread of epidemics; and to Mrs C.M. Best and Mrs J. Stuart for typing the thesis so efficiently.

TABLE OF CONTENTS

NOTATION

Symbol	Usual Meaning
$X_i(t)$	number of individuals in colony i at time $t \geq 0$
a_i	initial number of individuals in colony i at time $t = 0$
λ_i	birth rate
μ_i	death rate
ν_i	migration rate
α_i	immigration rate
ξ_i	$\lambda_i - \mu_i - \nu_i$
σ_i	$\lambda_i + \mu_i + \nu_i$
ω_1, ω_2	$\frac{1}{2}[(\xi_1 + \xi_2) \pm \{(\xi_1 - \xi_2)^2 + 4\nu_1\nu_2\}^{\frac{1}{2}}]$
$p_{ij}(t)$	probability that $X_1(t)=i$ and $X_2(t)=j$: two-colony model
$G(\underline{z};t)$	probability generating function of the $\{X_i(t)\}$
$K(\underline{\theta};t)$	cumulant generating function of the $\{X_i(t)\}$
z_i, θ_i	dummy variables
$g_j(\underline{z};t)$	probability generating function of the $\{X_i(t)\}$: $\alpha_1 = \alpha_2 = 0$ and $a_j = 1$, $a_i = 0$ $(i \neq j)$
$f_{jt}^{(n)}(\underline{z})$, $g_{jt}^{(n)}(\underline{z})$ and $h_{jt}^{(n)}(\underline{z})$	sequences of probability generating functions which converge to $g_j(\underline{z};t)$ as $n \to \infty$
$m_i(t)$	mean number of individuals in colony i at time $t \geq 0$
$\psi(t)$	$m_1(t)/(m_1(t)+m_2(t))$

Symbol	Usual Meaning
$m_i^{(j)}(t)$	expansion of $m_i(t)$ up to and including terms of order ν^j
$Q_i^{(j)}(\nu t)$	$m_i(t)/m_i^{(j)}(t)$
$V_{ij}(t)$	second-order moments of population size
$\rho(t)$	correlation coefficient : $\rho(t) = V_{12}(t)/\{V_{11}(t)V_{22}(t)\}^{\frac{1}{2}}$
π_{ij}	equilibrium probabilities : $\pi_{ij} = \lim_{t\to\infty} p_{ij}(t)$
$p_i(t)$	probability of extinction : $p_i(t) = g_i(\underline{0};t)$
q_i	probability of ultimate extinction : $q_i = \lim_{t\to\infty} p_i(t)$
$p(\underline{x};t)$	probability that $X_i(t) = x_i$ $(-\infty < i < \infty)$: stepping-stone model
ξ	$\lambda - \mu - 2\nu$: stepping-stone model
$\underline{e}_k$	$(\ldots,0,1,0,\ldots)$ denotes the infinite-dimensional vector with 1 in the k^{th} place and zeros elsewhere
$\{h_r\}$	migration distribution
$f(\underline{r};t)d\underline{r}$	mean number of individuals in the element $d\underline{r}$ at time t
c_+, c_-	asymptotic velocities : nearest-neighbour migration
v_+, v_-	asymptotic velocities : diffusion approximation
$*$	Laplace transformation (for example, $m_i^*(s) = \int_0^\infty e^{-st} m_i(t)dt$)
$\sim$	$a(t) \sim b(t)$ means that $a(t)/b(t) \to 1$ as $t \to \infty$

Symbol	Usual Meaning
δ_j^i	Kronecker delta function
$\delta(x)$	Dirac delta function
$(x)_n$	$x(x-1)\ldots(x-n+1)$
$s(n,k)$	Stirling numbers of the first kind
$\Gamma(x)$	Gamma function
$I_n(x)$	modified Bessel function of the first kind
$K(r)$	complete elliptic integral of the first kind
$E(r)$	complete elliptic integral of the second kind
${}_2F_1(a,b;c;z)$	Gauss hypergeometric series

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION TO THESIS

1.1 A Survey of Basic Birth-Death-Migration Models

The birth and death process was introduced by Feller (1939) to provide a mathematical description of the growth of a population which develops under the influence of the following simple laws of reproduction and mortality.

(i) The events that might occur to an individual in the small time interval $(t,t+\delta t)$ are independent of the events occurring to other individuals and of the events that occurred to this individual in the past.

(ii) The risks of mortality and reproduction are assumed to be the same for each member of the population.

(iii) An individual known to be alive at time t has probability $\lambda\delta t + o(\delta t)$ of giving birth (splitting into two) and probability $\mu\delta t + o(\delta t)$ of dying in the small time interval $(t,t+\delta t)$.

(iv) The parameters λ and μ are independent of the time t .

Let the population size at time t be denoted by $X(t)$, with

$$p_n(t) = \Pr[X(t) = n \mid X(0) = 1] \quad (n = 0,1,\ldots) .$$

Then under assumptions (i)-(iv) Feller showed that the probabilities $p_n(t)$, which completely determine the structure of the process, satisfy the set of differential-difference equations

$$dp_n(t)/dt = (n+1)\mu p_{n+1}(t) + (n-1)\lambda p_{n-1}(t) - n(\lambda+\mu)p_n(t) \quad (n = 1,2,\ldots)$$

(1.1)

$$dp_0(t)/dt = \mu p_1(t) .$$

Although Feller (1939) proved that the mean and variance of $X(t)$ are given by

$$E(X(t)) = e^{(\lambda-\mu)t}$$

and

$$\mathrm{Var}(X(t)) = \left[\frac{\lambda+\mu}{\lambda-\mu}\right] e^{(\lambda-\mu)t}(e^{(\lambda-\mu)t} - 1),$$

respectively, a solution to his equations (1.1) was first given by C. Palm in an unpublished letter to N. Arley. Palm's formulae can be written as

$$p_0(t) = a(t)$$

(1.2)

$$p_n(t) = (1-a(t))(1-b(t))(b(t))^{n-1} \qquad (n = 1,2,\ldots)$$

where

$$\frac{a(t)}{\mu} = \frac{b(t)}{\lambda} = \frac{e^{(\lambda-\mu)t} - 1}{\lambda e^{(\lambda-\mu)t} - \mu} .$$

An account of Palm's work is given by Arley and Borchsenius (1945).

The population whose growth is described by equations (1.1) is rather far removed from reality, and a great deal of attention has been paid to processes for which assumptions (i)-(iv) have been relaxed or altered in some way. In particular, Kendall (1948a) gives the complete solution of the equations governing the generalized birth-death process in which the birth and death rates $\lambda(t)$ and $\mu(t)$ may be any specified functions of the time t ; the most interesting example is perhaps that in which $\lambda(t)$ and $\mu(t)$ are periodic functions of t . A thorough study of 'general' birth and death processes has been carried out by Karlin and Mc.Gregor (1957a, 1957b).

Kendall (1948b) goes on to relax the Markov property (iii). This implies, for example, that in the absence of mortality an individual born at time t will itself undergo subdivision at a time $t+s$, where the 'generation time' s has the exponential distribution

$$\lambda e^{-\lambda s} \qquad (s \geq 0) \ .$$

Kendall remarks that this is very different from the distributions of generation time often observed in practice, and he suggests a modified process in which s is distributed as a χ^2_{2k} - variate, where k is an integer greater than unity. When a new individual is born, it passes through a series of phases, k in number, and only after it has attained the k^{th} phase can it undergo subdivision. The lifetime in each phase is assumed to have the exponential distribution

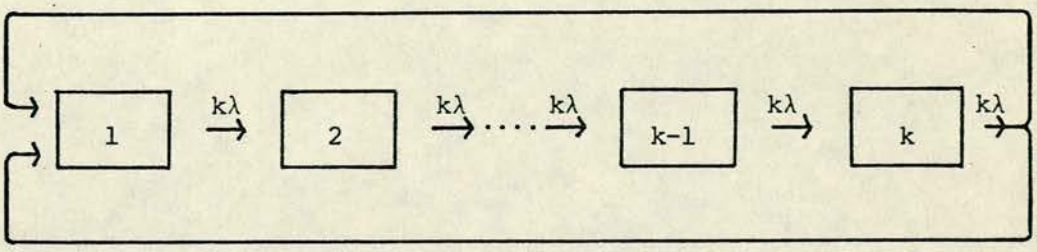
$$k\lambda e^{-k\lambda T} \quad (T \geq 0) \ ,$$

different lifetimes being independent. This multi-phase birth process possesses the Markov property, provided that its development is described by the vector variate $\underline{X}(t) = (X_1(t), \ldots, X_k(t))$, where $X_i(t)$ $(1 \leq i \leq k)$ denotes the number of individuals in the i^{th} phase at time t . Kendall's analysis is performed for the purely reproductive process $(\mu=0)$.

Kendall's multi-phase process, illustrated in figure 1.1, can be envisaged as representing a population which is spatially distributed amongst k sites or colonies. Individuals in colony i $(1 \leq i < k)$ may migrate to colony $i+1$, whilst in the event of an individual in

Figure 1.1

Kendall's multi-phase process represented as a population which is spatially distributed amongst k colonies



colony k giving birth, both it and its offspring instantaneously migrate to colony 1. This thesis is devoted to the mathematical analysis of the effect of migration between colonies, each of which is developing according to a simple birth-death-immigration process.

Consider a model based on the following assumptions.

(i) The habitat is divided into a number of colonies. Within a colony the species can be regarded as a single population, spatial separation being ignored. Between neighbouring colonies, migration is common, having an effect on the population size in a colony of the same order of magnitude as reproduction, immigration and death. Thus the model can represent either a continuous habitat, divided arbitrarily into colonies for mathematical description, or a discontinuous habitat in which the discontinuities are not serious barriers to movement.

(ii) The effect of migration is immediate. That is, the time taken to migrate is small compared with the life cycle.

(iii) Migration is 'conservative'; that is, there are no losses during migration, so that an individual leaving one colony arrives at another.

[Emigration out of the system may be considered as being equivalent to 'death', and shall therefore not be explicitly considered in this thesis.]

We shall regard this model as being a special type of Markov population process, these processes having frequently been used to represent situations involving numbers of individuals in different categories or colonies. In such processes the state at any time t is represented by the vector $\underline{X}(t) = (X_1(t),\ldots,X_k(t))$, where $X_i(t)$ denotes the number of individuals in the i^{th} colony at time t,

and the random evolution of $\underline{X}(t)$ is that of a continuous-time Markov chain. The jumps of the chain may be of three types, corresponding to the arrival of a new individual, the departure of an existing one, or the transfer of an individual from one colony to another. Kingman (1969) presents a systematic account of the methods available to analyse these processes, with particular reference to the calculation of stationary distributions (when they exist). The methods described yield solutions for quite a large class of interesting processes.

Kingman denotes the transition rates by

$$q(\underline{n},\underline{n}+\underline{e}_i) = \alpha_i(\underline{n})$$

$$q(\underline{n},\underline{n}-\underline{e}_i) = \beta_i(\underline{n})$$

$$q(\underline{n},\underline{n}-\underline{e}_i+\underline{e}_j) = \gamma_{ij}(\underline{n}) \qquad (i \neq j) ,$$

where $\underline{e}_i$ is the vector with all components zero except for 1 in the i^{th} place, and he remarks that in many examples of Markov population processes the arrival, departure and transfer rates depend only on the numbers in the colonies affected by the transition. The process is described as <u>simple</u> if it is possible to write

$$\alpha_i(\underline{n}) = \alpha_i(n_i) \quad ; \quad \beta_i(\underline{n}) = \beta_i(n_i) \quad ; \quad \gamma_{ij}(\underline{n}) = \gamma_{ij}(n_i,n_j) .$$

The transition rates of the models developed in this thesis have the simpler form given by

$$(1.3) \quad \alpha_i(n_i) = \lambda_i n_i + \alpha_i \quad ; \quad \beta_i(n_i) = \mu_i n_i \quad ; \quad \gamma_{ij}(n_i,n_j) = \nu_{ij} n_i \quad .$$

The parameters λ_i, α_i, μ_i and ν_{ij} correspond to birth, immigration, death and migration, respectively.

Although the transition rates (1.3) are extremely simple in form, the mathematical analysis surrounding them is complex. To make these

rates reflect more realistic situations would almost certainly lead to mathematical intractability. For example, the growth of any population in a restricted environment must inevitably be limited by a shortage of resources, and a stochastic analysis of the corresponding single-colony logistic model alone involves an order of difficulty many times greater than that of the simple birth-death process. Also implicitly assumed is that the transition rates (1.3) are independent of age. However, it is known (Deevy (1947)) that for many species of birds the death rate is nearly independent of age once they have become adult, and the same is probably true of many species of fish. A further objection is that we are restricting our model to a single-species, single-sex population, although if the species is bisexual we may consider only the females and postulate that there is never a shortage of males. For a discussion of these and many other factors affecting population growth we refer the interested reader to Pielou (1969), May (1973) and Maynard Smith (1974). Arnason (1971) devotes the first half of his thesis to an investigation of the suitability of Markov processes for describing the migration of animal populations. The second part details an investigation of how the parameters of the models can be estimated.

In Chapters 2 to 7 I investigate the probabilities

$$p_{ij}(t) = \Pr[X_1(t)=i,\ X_2(t)=j \mid X_1(0)=a_1,\ X_2(0)=a_2] \quad (i,j=0,1,\ldots)$$

corresponding to the transition rates (1.3) for the two-colony process. Here a_1 and a_2 denote the initial population sizes of the two colonies. The Kolmogorov forward differential equation for the probability generating function (p.g.f.)

$$G(z_1,z_2;t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij}(t) z_1^i z_2^j$$

is developed in Chapter 2, and from it I derive the means, variances and covariances. Although this differential equation has in general been found impossible to solve directly, in the sense that at the present time we are unable to write down an explicit expression for its solution, exact solutions may be obtained for certain special cases in which some of the parameters are placed equal to zero. In Chapter 3 I obtain exact solutions for three such cases, and develop a recursive solution for the probabilities $p_{ij}(t)$ in a fourth.

The complexity of the solution in one of these special cases underlines the extreme difficulty of deriving an exact solution in general, and in Chapter 4 we turn our attention to the derivation of approximate stochastic solutions. First I approximate the process itself by modifying the birth mechanism, thereby developing an exact solution for this new process which closely relates to the Poisson distribution. Whilst the first-order moments of the original and modified processes coincide, the second-order moments do not. This defect is partly remedied by the fitting of a bivariate negative-binomial distribution. The chapter concludes with a discussion on approximations which are derived by placing an upper bound on the total population size.

The analysis of the function $G(z_1,z_2;t)$ in Chapter 5 involves the use of two different iterative techniques on the corresponding backward equations. In the first I assume that the migration rates ν_{12} and ν_{21} are equal, and I obtain a power series solution in terms of this common migration rate. In the second I show how to generate sequences of functions, defined recursively by means of integral equations, that converge to the required solution.

If either immigration rate α_i $(i = 1,2)$ is non-zero, both colonies may be empty with non-zero probability at any particular time but extinction is impossible. However, if $\alpha_1 = \alpha_2 = 0$, ultimate extinction becomes a definite possibility, and Chapter 6 is devoted to an analysis of the extinction probability $p_{00}(t)$.

In Chapter 7 simulation techniques are used to compute equilibrium probabilities $p_{ij}(\infty)$ and to obtain realizations of the process for suitable choices of parameter values. A particular simulation study is undertaken in order to illustrate the effect of the magnitude of the migration parameters ν_{12} and ν_{21} on the goodness-of-fit of the two approximate distributions developed in Chapter 4.

So far I have discussed the effect of migration between two separate geographic regions on population size, and although many of the results in Chapters 2 to 7 could (in theory) be generalized to cover migration between n separate regions $(n = 2,3,\ldots)$, little gain in understanding of the process would result from the exercise. However, if individuals may migrate between neighbouring colonies only, the situation becomes far more promising. In Chapter 8 I introduce the 'stepping-stone model' to examine the one-dimensional process, in which the population is composed of an infinite number of colonies which may be considered to be situated at the integer points i $(-\infty < i < \infty)$ of a single co-ordinate axis. In this model the vector $\underline{X}(t) = (X_1(t), X_2(t))$ is replaced by the infinite-dimensional vector $\underline{X}(t) = (\ldots, X_i(t), \ldots)$, where $X_i(t)$ denotes the number of individuals in colony i at time t, and an individual in colony i may migrate to colony j only if $j = i-1$ or $i+1$ (see section 8.1 for a complete description of the process).

The 'stepping-stone model' was originally proposed by Kimura (1953) to investigate the effect of the breeding structure of a population on genetic differentiation within the population. A textbook account of this associated genetic problem is to be found in section 9.9 of Crow and Kimura (1970) and further references are given in section 8.1 of this thesis.

As well as possessing strong connections with problems in genetics, this one-dimensional birth-death-migration model is also directly relevant to the spread of epidemics. For suppose that every colony consists solely of N susceptibles, and then let an infected individual enter colony 0 at time $t=0$ (say) with $X_i(t)$ $(-\infty < i < \infty)$ now denoting the number of infectives in colony i at time $t \geq 0$. I shall refrain from giving an account of the general theory of epidemics here, but instead refer the interested reader to Bailey (1957) and the excellent survey paper of Dietz (1967). Suppose that in the small time interval $(t,t+\delta t)$ an infective in colony i may infect a particular susceptible in the same colony with probability $\alpha\delta t + o(\delta t)$; or may be removed either by death or isolation with probability $\beta\delta t + o(\delta t)$; or may migrate either to colony $i+1$ or to colony $i-1$ with probability $\gamma\delta t + o(\delta t)$. Suppose further that susceptibles may not migrate. Then in the opening stages of the epidemic the vector process $\underline{X}(t)$ may be approximated by the above birth-death-migration process with parameter values

$$\lambda_i \doteq \alpha N \; ; \; \mu_i = \beta \; ; \; \nu_{i,i+1} = \nu_{i,i-1} = \gamma, \; \nu_{ij} = 0 \quad \text{(otherwise)}.$$

The analysis of the linear model with parameters $\lambda_i = \lambda$; $\mu_i = \mu$; $\nu_{i,i+1} = \nu_{i,i-1} = \nu$ and $\nu_{ij} = 0$ (otherwise) is given in Chapter 8. The Kolmogorov forward differential equation for the p.g.f.

corresponding to $\underline{X}(t)$ cannot be solved directly, and approximate stochastic solutions are developed in an analogous manner to those derived in Chapter 4 for the two-colony process. An exact stochastic solution is obtained for the special case in which births are excluded ($\lambda = 0$) and immigration is introducted from outside the system into colony 0. The first-order and second-order moments are derived for a more general version of this model in which $\lambda \geq 0$.

If the population has a positive rate of growth and is initially concentrated into a relatively small geographic region, we may expect it to diffuse into the surrounding areas and eventually to take over the entire territory. This expanding population may be envisaged as generating a travelling wave, and in Chapter 9 I investigate the velocity of propagation and the form of the wave profile corresponding to the stepping-stone model developed in Chapter 8. The velocities generated by this model are then compared with those obtained from a continuous state space approximation in which an individual is not constrained to lie at the integer points i $(-\infty < i < \infty)$ but may instead occupy any point x $(-\infty < x < \infty)$ on the real line. All these results are obtained by using deterministic arguments, and in order to assess the relevance of these deterministic results to their stochastic counterparts I conclude the chapter with a small simulation study.

Mathematical interest in the spatial propagation of population waves was aroused by two investigations in the 1930's into the advance of an advantageous gene along a linear habitat, such as a shore line. The theory subsequently developed includes the evolution of neutron populations, the spread of consumer demand, flame propagation, and the spread of epidemics (see section 9.1 for

references and a discussion of these and related processes). Whilst such processes may be regarded as being analogous to the nearest-neighbour migration models so far considered in this thesis, other phenomena such as the spread of stem rust and foot and mouth disease may not. These latter diseases may be transmitted by windborne spores, and so infection may travel a large distance from one site to another without infecting the areas in-between. In Chapter 10 I examine non-nearest-neighbour migration models and develop expressions for the mean sizes of each colony at time t for several appropriate migration distributions $\{\nu_{i,i+j} : -\infty < j < \infty\}$. Unfortunately, these results appear to be too complex to allow wavefront velocities to be determined from them.

In Chapters 2 to 10 I develop theoretical results for population models in which there is migration either between two colonies or between an infinite number of colonies situated at the integer points of a single co-ordinate axis. To conclude the thesis I present a spatial model in two-dimensions, and relate it to data on the spatial distribution of flour beetles in a closed container.

1.2 Numbering System

Figures and tables are numbered successively within a chapter, so that figure 3.4 (say) refers to the fourth figure in Chapter 3.

Equations are numbered successively within a section; for example, equation (4.3) in Chapter 2 refers to equation number 3 in section 4 of that Chapter. However, if the same equation were being referenced in Chapter 6 (say), it would carry the full label (2.4.3).

1.3 Published Results

With the permission of his supervisor, the author has published parts of this thesis in the following papers.

(1972) : §§2.3.iii; 9.7.i; 9.7.ii,

(1973a): §§2.5.i*; 2.5.iv
4.2.i*; 4.2.iii; 4.2.iv*
5.2.i; 5.2.iii; 5.3,

(1973b): §§2.3.i
3.2.ii; 3.2.iii; 3.3.i; 3.3.iv; 3.4.i
4.4.i,

(1974): §§8.1; 8.2.i; 8.2.iii; 8.3; 8.4.

*published in a more simplified form.

CHAPTER 2

TWO-COLONY BIRTH-DEATH-MIGRATION-IMMIGRATION MODELS

2.1 Examples of Related Processes

Two or more interconnected birth and death processes arise so often in practical situations that it is worthwhile illustrating the types of situation to which they relate. One such biological process that has recently attracted considerable attention in the literature is that of phage-reproduction.

Viruses are small particles of RNA (ribonucleic acid) or DNA (deoxy-ribonucleic acid) wrapped in a protein coat, which can be crystallized into a variety of regular, often polyhedral, shapes. They are much smaller than bacteria, and are capable of passage through filters designed to arrest these. Among the numerous viruses, bacteriophages (called phages for short) have been the subject of much concentrated study; the isolated DNA strand, used by phages as their genetic information carrier, makes them eminently suitable in investigations on the molecular basis of life.

When a phage infects a bacterium, its protein coating is left on the boundary, while a DNA strand enters the bacterium. After a brief period, the boundary of the bacterium alters so that no further phage can attach to it. The DNA strand (vegetative phage) multiplies within the bacterium by division, until a pool of about 40 to 80 phages has been reached; the period during which this occurs is referred to as the eclipse. After the eclipse phase (approximately 7 to 10 minutes), some vegetative phage mature,

i.e. receive a protein coating, and are no longer able to reproduce. This act of maturing may be considered as being equivalent to migration. When the infected bacterium bursts, about 200 to 300 mature phage are released, together with about 40 to 80 vegetative phage. The mature phage are able to attack other bacteria, whilst the vegetative phage cannot.

Models for phage reproduction within a bacterium have been discussed by many authors, and all are effectively birth-death type processes. For a detailed account of the subject the interested reader is recommended to study the excellent review paper by Gani (1965). Such birth-death type processes are directly relevant to the theory of genetics, with migration being interpreted as mutation. For example, the phage process may be used to represent the growth of a bacterial colony in which normal (a) and mutant (A) bacteria correspond to vegetative and mature phage, respectively. Once a bacterium mutates, it and its offspring remain of the mutant type and do not mutate back to the normal type. Armitage (1952) presents a review of early work describing the changes in a population of bacteria which are subject to birth and mutation.

Gani and Yeo (1965) chacterize this phage/genetic process by the following set of probabilities. They assume that in any small time interval $(t,t+\delta t)$ there are probabilities

$\lambda_1(t)\delta t + o(\delta t)$: of (a) splitting into (a,a) [birth],

$\lambda_2(t)\delta t + o(\delta t)$: of (A) splitting into (A,A) [birth],

$\mu_1(t)\delta t + o(\delta t)$: of (a) mutating to (A) [migration],

$\nu_1(t)\delta t + o(\delta t)$: of (a) dying,

$\nu_2(t)\delta t + o(\delta t)$: of (A) dying.

During eclipse we have $\nu_1(t) = \nu_2(t) = 0$ (their notation unfortunately switches the now established usage of μ and ν to represent death and migration, respectively). The authors discuss the probability density function (p.d.f.) of the number of survivors and deaths up to time t, and obtain an expression for the distribution of the time until the bacterium bursts.

In section 1.1 we noted that the opening stages of an epidemic which develops over the integer points i $(-\infty<i<\infty)$ may be approximated by an appropriate birth-death-migration model. The corresponding two-colony epidemic process may be treated in exactly the same way. Here we assume that the probability of the number of infectives $X_i(t)$ in colony i $(i=1,2)$ increasing by 1 in the small time interval $(t,t+\delta t)$, because a susceptible in colony i becomes infected, is approximately equal to

$$\alpha_i X_i(t) N_i \delta t + o(\delta t)$$

where N_i is the initial number of susceptibles in that colony at time $t=0$. Griffiths (1972) employs a variation of this approach to examine the effect of the introduction of disease into a previously uninfected area. In this case the susceptible populations of both hosts and vectors could be considered as approximately constant. If we now call infected hosts <u>type 1</u> individuals and infected vectors <u>type 2</u> individuals then the approximation reduces the process to a bivariate birth-death

process in which the birth rate of type 1 individuals is proportional to the total number of type 2 individuals and vice-versa. The death rate for each type is proportional to the total number of individuals of that type. He examines the behaviour of the process conditional upon extinction and analyses the distribution of cumulative population size.

2.2 The Kolmogorov Forward Equation

2.2.i Introduction

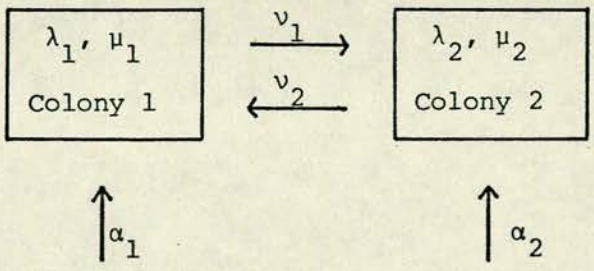
Consider a population divided into two colonies, and denote the number of individuals in colony i $(i=1,2)$ at time $t \geq 0$ by $X_i(t)$ with

$$X_i(0) = a_i \qquad (i=1,2) \ .$$

Next suppose that each colony i is subject to a simple stochastic birth-death-immigration-migration process with birth, death and immigration rates λ_i, μ_i and α_i respectively, and with migration rate ν_i to the other colony j. This situation is represented in figure 2.1. The immigration components α_1 and α_2 prevent the

Figure 2.1

Two-colony birth-death-migration-immigration process



population from becoming extinct and, under certain conditions, allow a limiting equilibrium distribution of population size to develop. Thus for $i = 1,2$ and $\delta t > 0$:

$\Pr[\text{an i-individual multiplies to two in } (t,t+\delta t)] = \lambda_i \delta t + o(\delta t)$;

$\Pr[\text{an i-individual dies in } (t,t+\delta t)] = \mu_i \delta t + o(\delta t)$;

$\Pr[\text{an i-individual transfers to colony j in } (t,t+\delta t)] = \nu_i \delta t + o(\delta t)$;

$\Pr[\text{an i-individual undergoes no transitions in } (t,t+\delta t)]$

$$= 1 - (\lambda_i + \mu_i + \nu_i)\delta t + o(\delta t);$$

$\Pr[\text{an i-individual undergoes more than one transition in } (t,t+\delta t)]$

$$= o(\delta t);$$

$\Pr[\text{an individual enters colony i from an outside system in } (t,t+\delta t)]$

$$= \alpha_i \delta t + o(\delta t).$$

Furthermore, it is assumed that all the events that might occur to an individual in the small time interval $(t,t+\delta t)$ are independent of the events which occur to other individuals and of the events that occurred to this individual in the past. With these assumptions the vector process $\{X_1(t), X_2(t)\}$ is a Markov process. Here the transition rates λ_i, μ_i, α_i and ν_i $(i=1,2)$ are non-negative and constant.

For $i,j = 0,1,2,\ldots,$ let the probabilities

$$p_{ij}(t) = \Pr[X_1(t)=i,\ X_2(t)=j \mid X_1(0)=a_1,\ X_2(0)=a_2] ,$$

and let

$$G(z_1,z_2;t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij}(t) z_1^i z_2^j \qquad (|z_i| \leq 1 ;\ i=1,2)$$

be their probability generating function. Consideration of all

possible events in the small time interval $(t,t+\delta t)$ yields the differential equations

$$\begin{aligned} dp_{ij}(t)/dt = {} & \lambda_1(i-1)p_{i-1,j}(t) + \lambda_2(j-1)p_{i,j-1}(t) + \mu_1(i+1)p_{i+1,j}(t) \\ & + \mu_2(j+1)p_{i,j+1}(t) + \nu_1(i+1)p_{i+1,j-1}(t) \\ & + \nu_2(j+1)p_{i-1,j+1}(t) + \alpha_1 p_{i-1,j}(t) + \alpha_2 p_{i,j-1}(t) \\ & - p_{ij}(t)[\, i(\lambda_1+\mu_1+\nu_1) + j(\lambda_2+\mu_2+\nu_2) + \alpha_1 + \alpha_2] \end{aligned} \tag{2.1}$$

where we define

$$p_{ij}(t) \equiv 0 \quad \text{for} \quad i<0 \quad \text{or} \quad j<0 \ .$$

The Kolmogorov forward differential equation for $G(z_1,z_2;t)$, obtained by multiplying both sides of equations (2.1) by $z_1^i z_2^j$ and summing the resultant expression over $i,j = 0,1,2,\ldots,$ is then given by

$$\begin{aligned} \frac{\partial G}{\partial t} = {} & [\,(\lambda_1 z_1 - \mu_1)(z_1-1) + \nu_1(z_2-z_1)]\frac{\partial G}{\partial z_1} \\ & + [\,(\lambda_2 z_2 - \mu_2)(z_2-1) + \nu_2(z_1-z_2)]\frac{\partial G}{\partial z_2} + [\alpha_1(z_1-1) + \alpha_2(z_2-1)]G \end{aligned} \tag{2.2}$$

with

$$G(z_1,z_2;0) = z_1^{a_1} z_2^{a_2} \ . \tag{2.3}$$

Bartlett ((1966), p.77-80) considers a similar process in which there are two types of particle (corresponding to individuals above), photons and electrons. These give rise respectively to pairs of electrons and photon-electron pairs. He derives the stochastic solution when the death rates are neglected and obtains differential equations for the first- and second-order factorial moments.

Although equation (2.2) has in general been found impossible to solve directly, certain results have been obtained for the special case when there is no immigration so that $\alpha_1 = \alpha_2 = 0$. Puri (1968) solves it for the case $\lambda_1 = \nu_2 = 0$, and obtains the solution to the corresponding Kolmogorov backward equations when $\lambda_2 = \nu_2 = 0$. Both Puri (1968) and Renshaw (1970, 1973a) develop sequences of functions, defined recursively by means of integral equations, that converge to the p.g.f.'s of these backward equations.

Approximate solutions may also be developed when immigration is present $(\alpha_1 + \alpha_2 > 0)$, as may exact solutions in several special cases when certain of the parameters equal zero. These approaches are discussed in Chapters 3 and 4. In a very recent paper Alksland (1975) considers the analogous n-colony process. He derives a solution in terms of a sequence of functions which converges to the required probability generating function, and he obtains exact solutions to the forward differential equation in two special cases.

Although the joint p.g.f. of $\{X_1(t), X_2(t)\}$ may not be found explicitly, the first- and second-order moments may be derived without too much difficulty and their derivation is described later on in this chapter. Still further knowledge about the behaviour of the process may be gained by simulation, and this technique is described in Chapter 7. I shall now briefly present some simulated realizations in order that the reader may gain an intuitive feel for the process before embarking on the mathematical analysis.

The way in which a simple one-colony birth-death-immigration process, with parameters λ, μ, α, respectively, develops is dependent on the sign of the 'net growth rate' $\lambda-\mu$. For if $m(t)$ denotes the mean number of individuals present in the colony at time t, the deterministic equation for $m(t)$ is given by

$$dm(t)/dt = (\lambda-\mu)m(t) + \alpha$$

and this integrates to give

$$m(t) = m(0)e^{(\lambda-\mu)t} + \left(\frac{\alpha}{\lambda-\mu}\right)(e^{(\lambda-\mu)t} - 1) . \tag{2.4}$$

Thus if

(i) $\lambda<\mu$ $\quad m(t) \to \alpha/(\mu-\lambda)$ $\quad$ as $t \to \infty$

(ii) $\lambda=\mu$ $\quad m(t) = m(0) + \alpha t$

(iii) $\lambda>\mu$ $\quad \log(m(t)) \sim (\lambda-\mu)t$ $\quad$ as $t \to \infty$.

A similar situation exists for the bivariate process $\{X_1(t), X_2(t)\}$, only the criterion $\lambda-\mu$ is now replaced by

$$\omega_1 = \tfrac{1}{2}[(\xi_1+\xi_2) + \{(\xi_1-\xi_2)^2 + 4\nu_1\nu_2\}^{\frac{1}{2}}] \tag{2.5}$$

where $\xi_i = \lambda_i-\mu_i-\nu_i$ $(i=1,2)$ (this is proved in section 2.3). We shall therefore consider three separate simulations corresponding to $\omega_1<0$, $\omega_1=0$ and $\omega_1>0$.

2.2.ii A simulation example of two colonies with asymptotic equilibrium $(\omega_1<0)$

Figure 2.2a shows a realization of the process when $\omega_1 = -0.2$. The parameters chosen are as follows.

	λ_i	μ_i	ν_i	α_i	a_i
Colony 1	1.8	2.0	1.0	2.0	6
Colony 2	1.8	2.0	0.5	2.0	14

Here the birth, death and immigration parameters are the same for both colonies, but the rate of migration from colony 1 to colony 2 is double that from colony 2 to colony 1. The state of the process is recorded every 0.1 time units over the time interval $0 \leq t \leq 10$. It should be noted that the actual behaviour of the process is far more 'spiky' than is suggested by the figure. For 955 events occurred in the period $0 \leq t \leq 10$, giving an average of 9.55 events for each plot.

The two horizontal lines represent $m_1(\infty)$ and $m_2(\infty)$, the mean sizes of the two colonies in equilibrium. There is clearly considerable variation about these lines with migration causing some degree of correlation between $X_1(t)$ and $X_2(t)$. In fact it follows from results (3.24) and (4.21) - (4.23) of this chapter that in equilibrium the first- and second-order moments for each colony are as follows.

	Colony 1	Colony 2
Mean	7.0588	12.9411
Variance	32.3982	96.7940
Standard Deviation	5.6920	9.8384
Covariance	35.4031	

The linear correlation coefficient between $X_1(\infty)$ and $X_2(\infty)$ is 0.6322.

A cross-spectral analysis between $X_1(t)$ and $X_2(t)$ might well produce some interesting results, but we shall not pursue the time-series approach here.

2.2.iii A simulation example of two colonies with asymptotic linear growth ($\omega_1=0$)

If ω_1 takes any negative value the process fluctuates around the mean equilibrium values $m_1(\infty)$ and $m_2(\infty)$, whilst if ω_1 takes any positive value the process exhibits exponential growth. However, if ω_1 is zero the mean population sizes $m_1(t)$ and $m_2(t)$ ultimately have the linear form (see expression (3.25))

$$m_i(t) = c_i + d_i t \qquad (i=1,2) ,$$

for some constants c_i and d_i, and the process will fluctuate with ever increasing amplitude around this 'linear trend'.

To illustrate this situation figure 2.2b shows a realization of this process for the parameter values given below.

	λ_i	μ_i	ν_i	α_i	a_i
Colony 1	2.0	3.0	2.0	1.0	0
Colony 2	3.0	2.5	1.5	0.0	3

It follows from result (3.25) that for reasonably large t

$$m_1(t) \sim t/4$$
$$m_2(t) \sim t/2 \quad .$$

The state of the process is again recorded every 0.1 time units. Thus as the rate of occurrence of events is a linear combination of the two population sizes the number of events

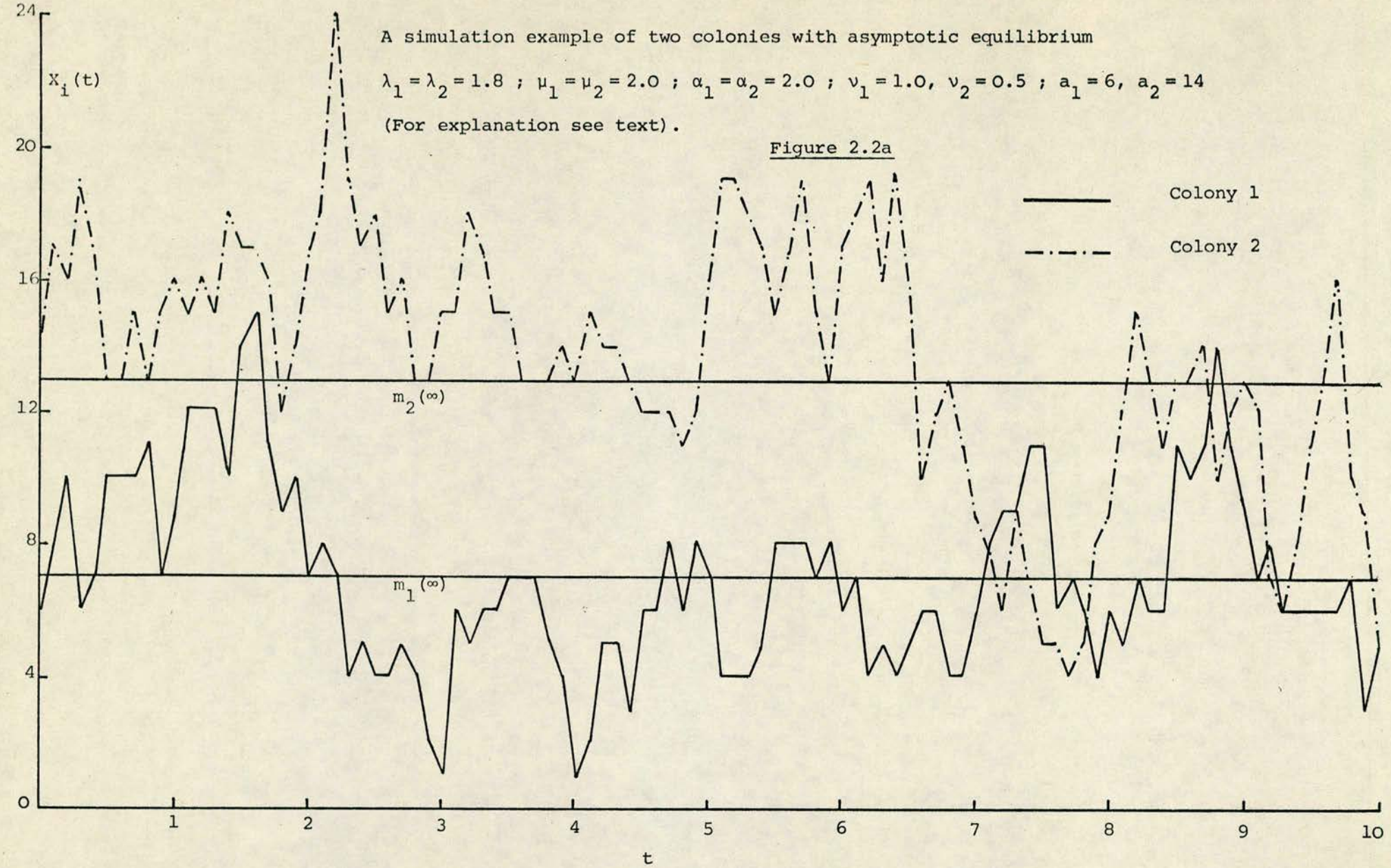

A simulation example of two colonies with asymptotic equilibrium
$\lambda_1 = \lambda_2 = 1.8$; $\mu_1 = \mu_2 = 2.0$; $\alpha_1 = \alpha_2 = 2.0$; $\nu_1 = 1.0$, $\nu_2 = 0.5$; $a_1 = 6$, $a_2 = 14$
(For explanation see text).
Figure 2.2a
Colony 1
Colony 2
$X_i(t)$
$m_2(\infty)$
$m_1(\infty)$
24
20
16
12
8
4
0
1
2
3
4
5
6
7
8
9
10
t

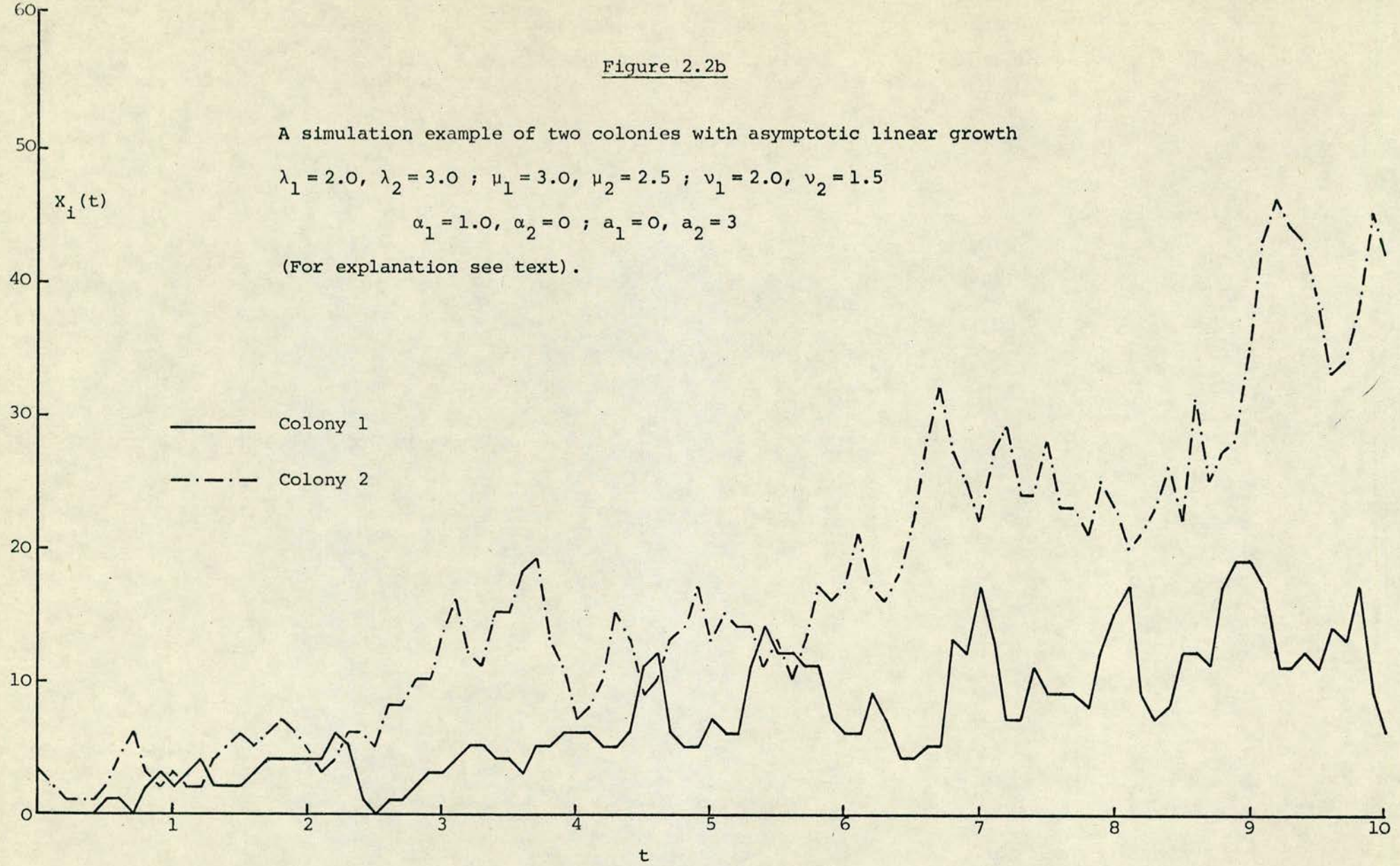
Figure 2.2b

A simulation example of two colonies with asymptotic linear growth

$\lambda_1 = 2.0$, $\lambda_2 = 3.0$; $\mu_1 = 3.0$, $\mu_2 = 2.5$; $\nu_1 = 2.0$, $\nu_2 = 1.5$

$\alpha_1 = 1.0$, $\alpha_2 = 0$; $a_1 = 0$, $a_2 = 3$

(For explanation see text).



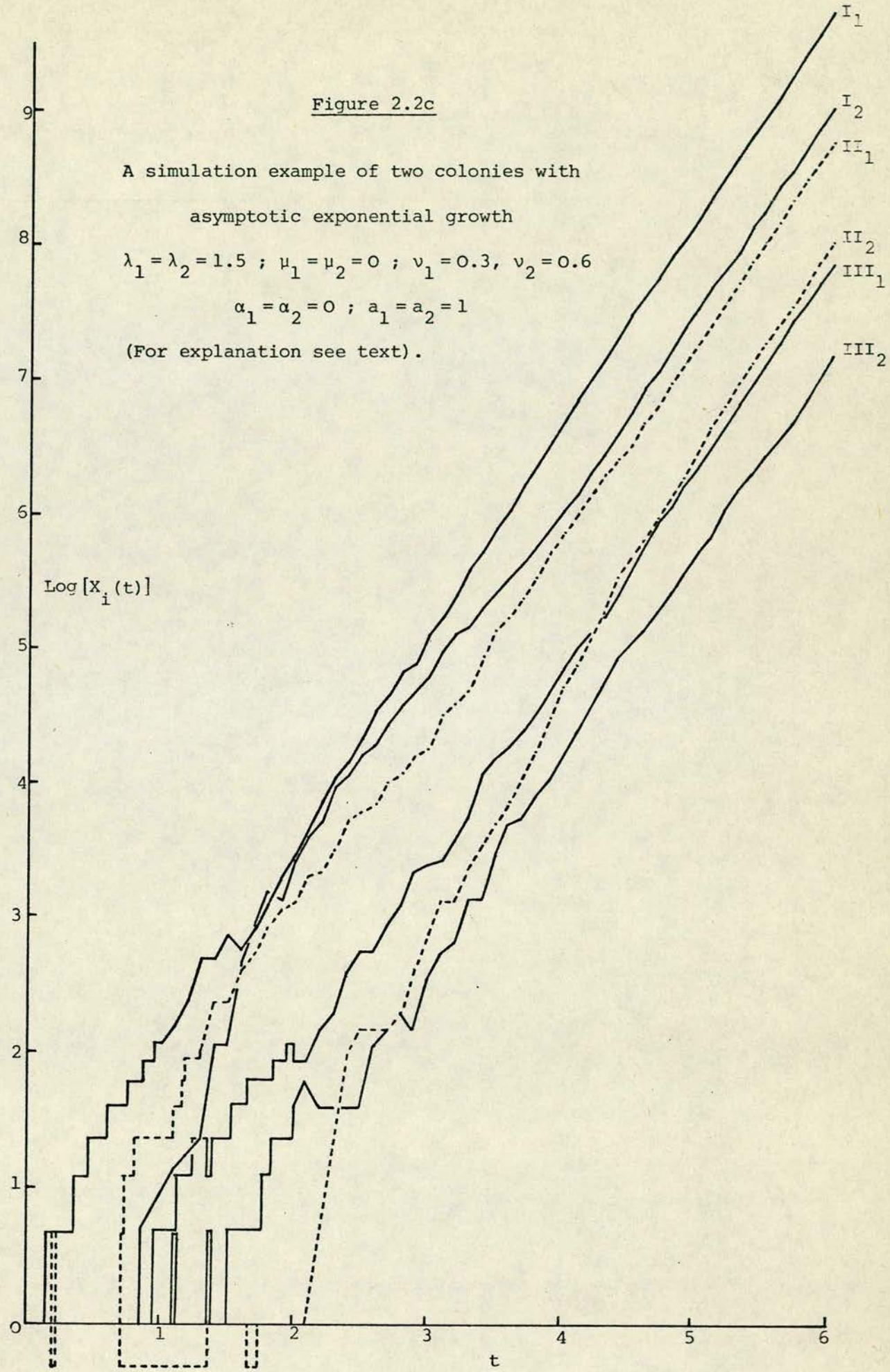


Figure 2.2c

A simulation example of two colonies with asymptotic exponential growth

$\lambda_1 = \lambda_2 = 1.5$; $\mu_1 = \mu_2 = 0$; $\nu_1 = 0.3$, $\nu_2 = 0.6$

$\alpha_1 = \alpha_2 = 0$; $a_1 = a_2 = 1$

(For explanation see text).

occurring in each successive time period of length 0.1 increases with t . For example, exactly 90 events occurred in each of the intervals $0 \leq t \leq 2.03$ and $10.60 \leq t \leq 10.70$. In the former

$$0 \leq X_1(t) \leq 7 \quad \text{and} \quad 1 \leq X_2(t) \leq 9 ,$$

whilst in the latter

$$21 \leq X_1(t) \leq 31 \quad \text{and} \quad 61 \leq X_2(t) \leq 78 .$$

2.2.iv A simulation example of two colonies with asymptotic exponential growth ($\omega_1 > 0$)

Figure 2.2c shows three realizations of the process when $\omega_1 = 1.5$, the parameter values used being as follows.

	λ_i	μ_i	ν_i	α_i	a_i
Colony 1	1.5	0	0.3	0	1
Colony 2	1.5	0	0.6	0	1

Thus the population develops as a pure birth process divided into two colonies, with the rate of migration from colony 2 to colony 1 double that from colony 1 to colony 2. Altogether 6 simulations were run and realizations I_i, II_i and III_i $(i=1,2)$ illustrated in figure 2.2c had the largest, third largest, and smallest total population size at time $t=6$. The subscripts $i=1$ and $i=2$ correspond to colonies 1 and 2 respectively in each case. Each successive event is shown for values of t less than t' where $1.0 \leq t' \leq 2.0$ (the value of t' depends on the particular realization being considered), whilst the state of the process is recorded every 0.1 time units for subsequent values. Note that in figure 2.2c the ordinate represents

the logarithm of population size, $\log X_i(t)$ $(i=1,2)$, and the state corresponding to $X_i(t)=0$ is represented by a line just below the abscissa.

It follows from the theory of branching processes that the vector random variable

$$(Y_1(t),\ Y_2(t)) = e^{-\omega_1 t}\,(X_1(t),\ X_2(t)) \tag{2.6}$$

converges almost surely to the vector random variable

$$(Y_1,\ Y_2) = \lim_{t\to\infty} (Y_1(t),\ Y_2(t))$$

as $t\to\infty$ (see, for example, Mode (1971)). Thus for large t

$$\log X_i(t) \sim \omega_1 t + \log Y_i \qquad (i=1,2)\ . \tag{2.7}$$

This suggests an asymptotically linear growth for $\log X_i(t)$ which can be considered as starting at the 'random time'

$$t_o = -(1/\omega_1)\log Y_i\ .$$

The variable t_o may be described as the 'stochastic lag' (Waugh (1972) discusses the corresponding univariate case). It is apparent from figure 2.2c that the stochastic lag for a particular realization depends on the development when t is small. The longer the time taken to reach a predetermined level, given by $X_i(t)=4$ (say), the larger the corresponding value of t_o.

For the particular case considered here we have the parameter values $\omega_1 = \lambda = 1.5$. It follows from (2.7) that provided t is sufficiently large the gradients of the 12 simulated curves, $\log X_i(t)$, should be centred around the value 1.5. A visual

inspection of the three simulations I_i, II_i and III_i suggests that the curves may be considered linear for $t \geq 5$. In fact the 12 values of

$$\tfrac{1}{2}[X_i(7) - X_i(5)] \qquad (i=1,2;\ 6 \text{ simulations})$$

have a range of (1.48, 1.56) and a mean value of 1.52. This is in close agreement with the expected value 1.5, if we remember that (2.7) is an asymptotic result and here t lies in the range 0 to 7.

2.3 Mean Population Sizes

2.3.i General solutions

Let $m_i(t)$ $(i=1,2)$ denote the mean number of individuals in colony i at time $t \geq 0$. Then on differentiating both sides of equation (2.2) with respect to z_i and placing $z_1 = z_2 = 1$ we get the first-order linear differential equations

$$(3.1) \quad dm_i(t)/dt = (\lambda_i - \mu_i - \nu_i)m_i(t) + \nu_j m_j(t) + \alpha_i \qquad (i,j=1,2;\ j \neq i)$$

where

$$(3.2) \qquad m_i(0) = a_i \qquad (i=1,2) .$$

These are clearly identical with the deterministic equations, as indeed would be expected because the individuals develop independently.

Let

$$\xi_i = \lambda_i - \mu_i - \nu_i \qquad (i=1,2)$$

and denote the Laplace transform of a function $f(t)$ by

$$(3.3) \qquad \mathcal{L}[f(t)] \equiv f^*(s) = \int_0^\infty e^{-st} f(t)\,dt \qquad , \Re(s) > 0 .$$

Then on applying the transformation (3.3) to equations (3.1) we have

$$(s-\xi_1)m_1^*(s) - \nu_2 m_2^*(s) = a_1 + \alpha_1 s^{-1} \tag{3.4}$$

$$(s-\xi_2)m_2^*(s) - \nu_1 m_1^*(s) = a_2 + \alpha_2 s^{-1} \tag{3.5}$$

for $\Re(s)$ large enough to ensure the existence of $m_1^*(s)$ and $m_2^*(s)$. Solving these two simultaneous equations for $m_1^*(s)$ and $m_2^*(s)$ we get

$$m_1^*(s) = [a_1 s + (\alpha_1 - a_1\xi_2 + \nu_2 a_2) + s^{-1}(\alpha_2\nu_2 - \alpha_1\xi_2)]/(s-\omega_1)(s-\omega_2) \tag{3.6}$$

$$m_2^*(s) = [a_2 s + (\alpha_2 - a_2\xi_1 + \nu_1 a_1) + s^{-1}(\alpha_1\nu_1 - \alpha_2\xi_1)]/(s-\omega_1)(s-\omega_2) \tag{3.7}$$

where ω_1 and ω_2 are the roots of the equation

$$\omega^2 - \omega(\xi_1+\xi_2) + (\xi_1\xi_2 - \nu_1\nu_2) = 0 , \tag{3.8}$$

namely

$$\omega_1, \omega_2 = \tfrac{1}{2}[(\xi_1+\xi_2) \pm \{(\xi_1-\xi_2)^2 + 4\nu_1\nu_2\}^{\frac{1}{2}}] . \tag{3.9}$$

Expressions (3.6) and (3.7) are inverted by first expanding their right hand sides into partial fractions and then applying the result

$$(s-a)^{-1} \equiv \mathcal{L}[e^{at}] . \tag{3.10}$$

If $\omega_1 \neq \omega_2$ this procedure yields

$$\begin{aligned} m_1(t) = (\omega_1-\omega_2)^{-1}\{ & e^{\omega_1 t}[a_1\omega_1 + \alpha_1 - a_1\xi_2 + \nu_2 a_2 + \omega_1^{-1}(\alpha_2\nu_2 - \alpha_1\xi_2)] \\ & - e^{\omega_2 t}[a_1\omega_2 + \alpha_1 - a_1\xi_2 + \nu_2 a_2 + \omega_2^{-1}(\alpha_2\nu_2 - \alpha_1\xi_2)]\} \\ & + (\omega_1\omega_2)^{-1}(\alpha_2\nu_2 - \alpha_1\xi_2) \end{aligned} \tag{3.11}$$

together with

$$m_2(t) = (\omega_1-\omega_2)^{-1}\{e^{\omega_1 t}[a_2\omega_1+\alpha_2-a_2\xi_1+\nu_1 a_1+\omega_1^{-1}(\alpha_1\nu_1-\alpha_2\xi_1)]$$

$$-e^{\omega_2 t}[a_2\omega_2+\alpha_2-a_2\xi_1+\nu_1 a_1+\omega_2^{-1}(\alpha_1\nu_1-\alpha_2\xi_1)]\} \tag{3.12}$$

$$+(\omega_1\omega_2)^{-1}(\alpha_1\nu_1-\alpha_2\xi_1) \quad .$$

If $\omega_1=\omega_2$ we see from (3.9) that $\xi_1=\xi_2=\xi$ (say) and $\nu_1\nu_2=0$. Let us therefore assume that $\nu_2=0$. Expressions (3.6) and (3.7) now invert to give

$$m_1(t) = a_1 e^{\xi t} + (\alpha_1/\xi)(e^{\xi t}-1) \tag{3.13}$$

and

$$m_2(t) = a_2 e^{\xi t} + (\alpha_2/\xi)(e^{\xi t}-1) + \nu_1 a_1 t e^{\xi t} \tag{3.14}$$

$$+ (\nu_1\alpha_1/\xi^2)[1-(1-\xi t)e^{\xi t}] \quad .$$

If $\nu_2=0$ and $\xi_1 > \xi_2$ (say), it follows from (3.9) that $\omega_1=\xi_1$ and $\omega_2=\xi_2$.

Expressions (3.11)-(3.14) were verifed by back-substitution into equations (3.1).

Note that if immigration is not present, so that $\alpha_1=\alpha_2=0$, expressions (3.11) and (3.12) reduce to

$$m_1(t) = (\omega_1-\omega_2)^{-1}\{e^{\omega_1 t}[a_1(\omega_1-\xi_2)+\nu_2 a_2] - e^{\omega_2 t}[a_1(\omega_2-\xi_2)+\nu_2 a_2]\} \tag{3.15}$$

$$m_2(t) = (\omega_1-\omega_2)^{-1}\{e^{\omega_1 t}[a_2(\omega_1-\xi_1)+\nu_1 a_1] - e^{\omega_2 t}[a_2(\omega_2-\xi_1)+\nu_1 a_1]\} \quad . \tag{3.16}$$

Expressions (3.15) and (3.16) have also been derived by the use of standard matrix techniques as described, for example, in Bartlett (1966). (See Renshaw (1970, 1972)).

If the transition rates are time-dependent the form of equations (3.1) remains unchanged but the nature of their general solution will not be simple. However, if migration may occur in only one direction, for example if $\nu_2(t) \equiv 0$, equations (3.1) become

$$dm_1(t)/dt = \xi_1(t)m_1(t) + \alpha_1(t) \tag{3.17}$$

$$dm_2(t)/dt = \xi_2(t)m_2(t) + \nu_1(t)m_1(t) + \alpha_2(t) \tag{3.18}$$

and these integrate directly to give

$$m_1(t) = \exp\{\int_o^t \xi_1(s)ds\}[a_1 + \int_o^t \alpha_1(s)\exp\{-\int_o^s \xi_1(x)dx\}ds] \tag{3.19}$$

$$m_2(t) = \exp\{\int_o^t \xi_2(s)ds\}[a_2 + \int_o^t [\nu_1(s)m_1(s) + \alpha_2(s)] \times \exp\{-\int_o^s \xi_2(x)dx\}ds] \quad . \tag{3.20}$$

In particular, if we place $\alpha_i(t) \equiv 0$ $(i=1,2)$

$$m_1(t) = a_1\exp\{\int_o^t \xi_1(s)ds\} \tag{3.21}$$

$$m_2(t) = a_2\exp\{\int_o^t \xi_2(s)ds\} + a_1\int_o^t \nu_1(s)\exp\{\int_o^s \xi_1(x)dx + \int_s^t \xi_2(x)dx\}ds \quad . \tag{3.22}$$

Expressions similar to (3.21) and (3.22) are contained in Gani and Yeo (1965) and Puri (1968).

2.3.ii Asymptotic moment formulae

As $\omega_1 \geq \omega_2$ it is the sign of ω_1 that determines whether $m_1(t)$ and $m_2(t)$ increase asymptotically to infinity, or remain bounded, as $t \to \infty$. Thus on letting $t \to \infty$ in expressions (3.11) and (3.12) we see that for $\omega_1 \neq \omega_2$ and $i,j=1,2;\ i \neq j$

(i) $\omega_1 > 0$: $m_i(t) \sim (\omega_1 - \omega_2)^{-1} e^{\omega_1 t}[a_i\omega_1 + \alpha_i - a_i\xi_j + \nu_j a_j + \omega_1^{-1}(\alpha_j\nu_j - \alpha_i\xi_j)]$ (3.23)

(ii) $\omega_1 < 0$: $m_i(t) \to (\alpha_j\nu_j-\alpha_i\xi_j)/(\omega_1\omega_2)$

(3.24) $= (\alpha_j\nu_j-\alpha_i\xi_j)/(\xi_1\xi_2-\nu_1\nu_2)$

(iii) $\omega_1=0$: $m_i(t)\sim-\omega_2^{-1}(\alpha_i-a_i\xi_j+a_j\nu_j)+(\alpha_j\nu_j-\alpha_i\xi_j)\lim_{\omega_1\to 0}\dfrac{\omega_2 e^{\omega_1 t}+(\omega_1-\omega_2)}{\omega_1\omega_2(\omega_1-\omega_2)}$.

Now the limit in this last expression is

$$-(1+\omega_2 t)/\omega_2^2$$

and so

(3.25) $$m_i(t)\sim -t(\alpha_j\nu_j-\alpha_i\xi_j)/(\xi_1+\xi_2) \quad .$$

We may obtain similar results if $\omega_1=\omega_2$. For it follows from (3.13) and (3.14) that

(i) if $\xi>0$: then $m_1(t) \sim (a_1+\alpha_1/\xi)e^{\xi t}$

(3.26) $$m_2(t) \sim \nu_1 t e^{\xi t}(a_1+\alpha_1/\xi)$$

and so

$$m_2(t)/m_1(t) \sim \nu_1 t \quad ;$$

(ii) if $\xi<0$: then $m_1(t)\to -\alpha_1/\xi$

(3.27) $$m_2(t)\to (\nu_1\alpha_1-\xi\alpha_2)/\xi^2$$

and so both means tend to finite non-zero limits;

(iii) and if $\xi=0$: then taking the limit as $\xi\to 0$ in (3.13) and (3.14) we get

(3.28)
$$m_1(t) = a_1 + \alpha_1 t$$
$$m_2(t) = a_2 + (\alpha_2+\nu_1 a_1)t + \tfrac{1}{2}\nu_1\alpha_1 t^2 \quad .$$

This last result is as would be expected. For if we assume that $\nu_2=0$ (as we did before in section (2.3.i)), the condition

$\xi=0$ implies that the growth rates $\lambda_1-\mu_1-\nu_1$ and $\lambda_2-\mu_2$ are both zero. Thus (3.28) is equivalent to writing

$m_1(t)$ = initial number + mean number of immigrants in time t

$m_2(t)$ = initial number + mean number of immigrants in time t

+ [migration rate (ν_1)] $\times \int_0^t m_1(s)ds$.

2.3.iii Asymptotic proportions of mean population sizes

If we examine the relative proportions of $m_1(t)$ and $m_2(t)$ by defining

(3.29) $$\psi(t) = m_1(t)/(m_1(t)+m_2(t)) ,$$

we see from (3.23), (3.24) and (3.25) that for $\omega_1\neq\omega_2$ $\psi(t)$ will in general approach two different finite non-zero limits as $t\to\infty$ depending on whether $\omega_1>0$ or $\omega_1\leq 0$. In particular, if $\omega_1\leq 0$

(3.30) $$\psi(t)\to \frac{\alpha_2\nu_2-\alpha_1\xi_2}{\alpha_1(\nu_1-\xi_2)+\alpha_2(\nu_2-\xi_1)} \qquad (\omega_1\leq 0, \omega_1\neq\omega_2) .$$

Note that this result is independent of the initial colony sizes (a_1,a_2) . For with probability one all of the a_1+a_2 original members and their descendents will eventually die, and the population will then consist entirely of immigrants and their progeny.

However, if $\omega_1=\omega_2$ the situation changes. For as we have shown previously, the condition $\omega_1=\omega_2$ implies that $\nu_2=0$ and $\xi=\lambda_1-\mu_1-\nu_1=\lambda_2-\mu_2$. If $\xi>0$ both colonies have the same exponential 'growth rate', but migration from colony 1 to colony 2 creates an ever-increasing imbalance in their relative sizes. In fact, expressions (3.26), (3.27) and (3.28) yield

$$\psi(t) \sim \begin{cases} 1/\nu_1 t & (\xi > 0) \\ 2/\nu_1 t & (\xi = 0) \\ -\alpha_1\xi/[\alpha_1(\nu_1-\xi)-\alpha_2\xi] & (\xi < 0) \end{cases} . \tag{3.31}$$

When $\xi<0$ $\psi(t)$ has the same limit as given in (3.30).

The function $\psi(t)$ also tends to a finite non-zero limit when there is no immigration $(\alpha_1=\alpha_2=0)$. For denote

$$\omega_1-\xi_1 = c_1\nu_1 \quad ; \quad \omega_2-\xi_2 = c_2\nu_2 .$$

Then it follows from (3.15) and (3.16) that $\psi(t)$ may be written in the form

$$\psi(t) = \frac{(a_1+a_2c_1)e^{\omega_1 t} - c_1(a_2+a_1c_2)e^{\omega_2 t}}{(a_1+a_2c_1)(1-c_2)e^{\omega_1 t} + (a_2+a_1c_2)(1-c_1)e^{\omega_2 t}} \qquad (\omega_1\neq\omega_2)$$

which gives

$$\psi(t) = (1-c_2)^{-1} + O(e^{-(\omega_1-\omega_2)t}) . \tag{3.32}$$

This holds even if $\omega_1<0$, when the population becomes extinct with probability one.

Now

$$\begin{aligned} c_2 &= (\omega_2-\xi_2)/\nu_2 \\ &= [(\xi_1-\xi_2)-\{(\xi_1-\xi_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}}]/2\nu_2 \\ &< 0 \end{aligned}$$

(provided $\nu_1\nu_2>0$ whenever $\xi_1\geq\xi_2$). Hence we may choose positive integers a_1 and a_2 to 'almost' satisfy

$$a_2+a_1c_2 = 0 , \tag{3.33}$$

whence

$$\psi(t) = (1-c_2)^{-1} \qquad \text{for all } t \geq 0 ,$$

and in this sense we may call an initial distribution (a_1,a_2) which satisfies (3.33) stationary.

These results are better illustrated by a consideration of the special case $\lambda_1=\lambda_2=\lambda$, $\mu_1=\mu_2=\mu$ and $\alpha_1=\alpha_2=0$. Expressions (3.15) and (3.16) give

$$m_1(t) = e^{(\lambda-\mu)t}(\nu_1+\nu_2)^{-1}\{\nu_2(a_1+a_2) + (a_1\nu_1-a_2\nu_2)e^{-(\nu_1+\nu_2)t}\}$$

(3.34)

$$m_2(t) = e^{(\lambda-\mu)t}(\nu_1+\nu_2)^{-1}\{\nu_1(a_1+a_2) + (a_2\nu_2-a_1\nu_1)e^{-(\nu_1+\nu_2)t}\}.$$

Hence

$$m_1(t) + m_2(t) = (a_1+a_2)e^{(\lambda-\mu)t} ,$$

as would be expected, and

$$\psi(t) = \frac{\nu_2}{\nu_1+\nu_2} + \frac{(a_1\nu_1-a_2\nu_2)e^{-(\nu_1+\nu_2)t}}{(a_1+a_2)(\nu_1+\nu_2)} \tag{3.35}$$

which yields

$$\psi(\infty) = \nu_2/(\nu_1+\nu_2) .$$

Thus the proportion $\psi(t)$ converges to the proportion of the migration rates to each of the colonies. We see from (3.34) that for the stationary distribution (a_1,a_2) which satisfies

$$a_1\nu_1 = a_2\nu_2$$

we have

$$(1+\nu_1/\nu_2)m_1(t) = (1+\nu_2/\nu_1)m_2(t) = (a_1+a_2)e^{(\lambda-\mu)t} .$$

As an illustration of result (3.35) consider the simulation results for $\omega_1 > 0$ described in section 2.2.iv. Here the parameter values are given by $\lambda_i=1.5$; $\mu_i=\alpha_i=0$; $a_i=1$ $(i=1,2)$ with $\nu_1=0.3$ and $\nu_2=0.6$. At time $t=7$ the six simulated values of the random variable

$$X_1(t)/(X_1(t)+X_2(t))$$

have a range of (0.6643, 0.6804), and this small interval contains the value

$$\psi(7) = 0.6664$$

computed from expression (3.35).

2.3.iv Critical migration

If immigration does not occur $(\alpha_1=\alpha_2=0)$, the total mean population size $m_1(t)+m_2(t)$ will increase to infinity or decrease to zero, as t tends to infinity, according as $\omega_1>0$ or $\omega_1<0$, respectively.

Let us assume that both $\nu_1>0$ and $\nu_2>0$, so that migration may occur in either direction. On rewriting (3.9) as

$$\omega_1,\omega_2 = \tfrac{1}{2}[(\xi_1+\xi_2)\pm\{(\xi_1+\xi_2)^2+4(\nu_1\nu_2-\xi_1\xi_2)\}^{\frac{1}{2}}] \quad , \tag{3.36}$$

we see that $\omega_1=0$ if and only if

$$(\xi_1+\xi_2)\leq 0 \quad \text{and} \quad \xi_1\xi_2=\nu_1\nu_2 \quad . \tag{3.37}$$

Hence, under the assumption that $\nu_1\nu_2>0$, expressions (3.37) imply that $\xi_1<0$ and $\xi_2<0$. It therefore follows from (3.36) and (3.37)

that

$$\text{if } \xi_1\xi_2 > \nu_1\nu_2 \quad (\xi_1,\xi_2<0) \quad \text{then the population decreases,}$$

$$\text{if } \xi_1\xi_2 = \nu_1\nu_2 \quad (\xi_1,\xi_2<0) \quad \text{then the population remains constant,} \tag{3.38}$$

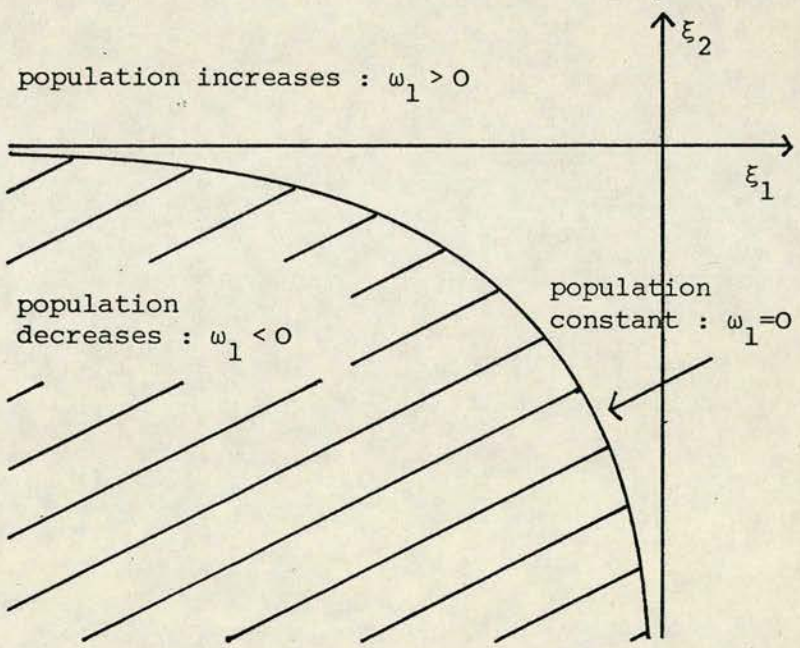
and the population increases in any other situation.

These results are illustrated in figure 2.3.

Figure 2.3

The critical migration curve as a function of ξ_1 and ξ_2

(For explanation see text)



It follows from this figure that if $\xi_1>0$ and $\xi_2<0$ (say), then no matter how large $-\xi_2$ becomes the total mean population size will still tend to infinity as $t\to\infty$. For the increase in colony 1 due to λ_1 will more than compensate for losses due to

μ_1 and ν_1, and so $m_1(t)$ will tend to infinity regardless of the magnitude of ξ_2.

If $\lambda_1-\mu_1<0$ and $\lambda_2-\mu_2>0$, an increase or decrease in the total population size will depend purely on the values of the migration parameters ν_i. If the ν_i are such that the total mean population size tends to a finite non-zero constant we shall say that we have a state of <u>critical migration</u>. Now we have seen from (3.37) that $\omega_1=0$ if and only if

$$\xi_1\xi_2 = \nu_1\nu_2 \quad ; \quad \xi_1<0\,,\ \xi_2<0 \quad . \tag{3.39}$$

Thus on denoting

$$d_i = \lambda_i-\mu_i \qquad (i=1,2)$$

we have

$$(d_1-\nu_1)(d_2-\nu_2) = \nu_1\nu_2 \quad ; \quad d_i-\nu_i<0 \quad (i=1,2)$$

whence

$$d_1d_2 = \nu_1d_2 + \nu_2d_1 \quad .$$

Suppose $d_1<0$ and $d_2>0$. Then, if $\nu_1,\nu_2>0$,

$$d_1-\nu_1 = \nu_2(d_1/d_2) < 0$$

and

$$d_2-\nu_2 = \nu_1(d_2/d_1) < 0 \quad .$$

Thus a state of critical migration will exist if and only if

$$d_1d_2 = \nu_1d_2 + \nu_2d_1 \quad ; \quad d_i<0,\ d_j>0 \quad (i\neq j) \quad . \tag{3.40}$$

If $d_1<0$ and $d_2>0$, then condition (3.40) determines the critical migration values as the set $\{\nu_1,\nu_2\}$ which satisfies

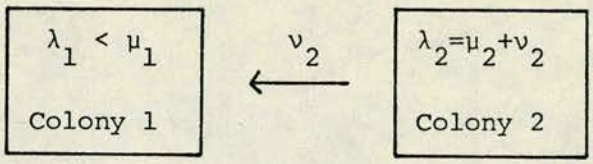
$$\nu_1 = d_1 - \nu_2(d_1/d_2) \quad ; \quad \nu_2>d_2 \quad . \tag{3.41}$$

We see from expression (3.41) that if $d_1<0$ and $d_2>0$ a necessary condition for critical migration to occur is $\nu_2>d_2$. Let us now examine what happens if $\nu_2=d_2$, so that $\xi_2=0$. It follows from (3.9) that

$$\omega_1 = \tfrac{1}{2}[\xi_1 + (\xi_1^2+4\nu_1\nu_2)^{\frac{1}{2}}] \quad ,$$

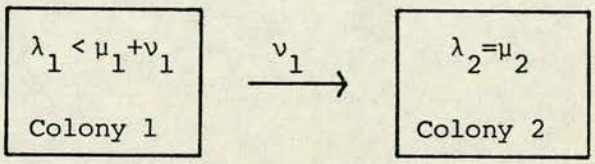
and so provided $\nu_1\nu_2=0$, we have $\omega_2<\omega_1=0$ for any $\xi_1<0$. This is easily explained by considering the following two situations.

(i) $\nu_1=0$:



Colony 2 has zero growth rate and migrants from colony 2 to colony 1 decay at rate $\mu_1 - \lambda_1$.

(ii) $\nu_2=0$:



Colony 1 eventually becomes empty; when this occurs colony 2 has a zero rate of growth.

So far we have considered $\omega_1\neq\omega_2$. If $\omega_1=\omega_2=0$, it follows from (3.36) that $\xi_1=\xi_2=0$ and $\nu_1\nu_2=0$. Take $\nu_2=0$. Adding expressions (3.28) we get

$$m_1(t) + m_2(t) = (a_1+a_2) + (\alpha_1+\alpha_2+\nu_1 a_1)t + \tfrac{1}{2}\nu_1\alpha_1 t^2 \quad ,$$

and so for critical migration to occur we must have $\alpha_1=\alpha_2=0$

together with either

(i) $\nu_1=0$: both colonies are completely separated from each other, and each behaves as a simple birth-death process with $\lambda_1=\mu_1$ and $\lambda_2=\mu_2$, or

(ii) $a_1=0$: colony 1 is permanently empty, and so colony 2 behaves as a simple birth-death process with $\lambda_2=\mu_2$.

2.4 Second-Order Moments of Population Size

2.4.i General solutions

The second-order moments may be examined in a manner similar to that used for the first. Let us put $z_i=\exp(\theta_i)$ $(i=1,2)$ and write the cumulant generating function (c.g.f.) as

$$K(\theta_1,\theta_2) = \log G(z_1,z_2;t) \quad .$$

Equation (2.2) then becomes

$$\begin{aligned}\frac{\partial K}{\partial t} = &[\lambda_1(e^{\theta_1}-1)+\mu_1(e^{-\theta_1}-1)+\nu_1(e^{\theta_2-\theta_1}-1)]\frac{\partial K}{\partial \theta_1} + \alpha_1(e^{\theta_1}-1) \\ &+ [\lambda_2(e^{\theta_2}-1)+\mu_2(e^{-\theta_2}-1)+\nu_2(e^{\theta_1-\theta_2}-1)]\frac{\partial K}{\partial \theta_2} + \alpha_2(e^{\theta_2}-1)\end{aligned} \tag{4.1}$$

where

$$K(\theta_1,\theta_2;0) = a_1\theta_1 + a_2\theta_2 \quad . \tag{4.2}$$

Let $V_{ii}(t)$ $(i=1,2)$ denote the variance of $X_i(t)$, and let $V_{ij}(t)$ $(i\neq j)$ denote the covariance of $X_i(t)$ and $X_j(t)$. Then on differentiating both sides of equation (4.1) twice with respect to θ_1 and θ_2 and placing $\theta_1=\theta_2=0$ we obtain the three first-order linear differential equations

$$(4.3)\quad dV_{11}(t)/dt = (\lambda_1+\mu_1+\nu_1)m_1(t) + \nu_2 m_2(t) + 2(\lambda_1-\mu_1-\nu_1)V_{11}(t) + 2\nu_2 V_{12}(t) + \alpha_1$$

$$(4.4)\quad dV_{12}(t)/dt = -\nu_1 m_1(t) - \nu_2 m_2(t) + \nu_1 V_{11}(t) + \nu_2 V_{22}(t) + [(\lambda_1-\mu_1-\nu_1)+(\lambda_2-\mu_2-\nu_2)]V_{12}(t)$$

$$(4.5)\quad dV_{22}(t)/dt = (\lambda_2+\mu_2+\nu_2)m_2(t) + \nu_1 m_1(t) + 2(\lambda_2-\mu_2-\nu_2)V_{22}(t) + 2\nu_1 V_{12}(t) + \alpha_2$$

where $V_{12}(t) \equiv V_{21}(t)$ for all $t \geq 0$.

Similar equations have previously been derived for the process with no immigration $(\alpha_1=\alpha_2=0)$ and the following solution has been suggested. Define the diagonal matrices

$$M(t) = \text{diag}\{m_i(t)\} \quad ; \quad H(t) = \text{diag}\{\xi_i m_i(t)+\nu_j m_j(t)\} \quad ; \quad D = \text{diag}\{\lambda_i\}$$

$(i,j=1,2:\ i\neq j)$ with

$$L = \begin{bmatrix} \xi_1 & \nu_1 \\ \nu_2 & \xi_2 \end{bmatrix} \quad \text{and} \quad V(t) = \begin{bmatrix} V_{11}(t) & V_{12}(t) \\ V_{21}(t) & V_{22}(t) \end{bmatrix}$$

Equations (4.3) - (4.5) may now be written in matrix form and integrated to yield the expression

$$(4.6)\qquad V(t) = \int_0^t e^{(t-s)L'} F(s)\, e^{(t-s)L}\, ds$$

for the variance-covariance matrix $V(t)$ where

$$F = H + M(D-L) + (D-L')M \quad .$$

(See Renshaw (1970,1972)). Note that in this vector-matrix context a dash (') denotes a transpose. We shall return to this approach in section 2.5.

It is worthwhile noting that our migration model is a special case of a multitype branching process and certain results may be obtained by reference to the general theory. For example, Harris ((1963), chapter 2, equation (4.3)) gives the expression corresponding to (4.6) for the multitype Galton-Watson process, and Mode ((1971), p.300-307) discusses the covariance function of the general multitype branching process in continuous time. An investigation of limiting properties and age-dependence using the theory developed by Mode (1971) might prove fruitful (see Crump (1970)).

Applying the Laplace transformation (3.3) to equations (4.3) - (4.5) we have

$$(4.7)\quad sV_{11}^{*}(s) = (\lambda_1+\mu_1+\nu_1)m_1^{*}(s) + \nu_2 m_2^{*}(s) + 2\xi_1 V_{11}^{*}(s) + 2\nu_2 V_{12}^{*}(s) + \alpha_1 s^{-1}$$

$$(4.8)\quad sV_{12}^{*}(s) = -\nu_1 m_1^{*}(s) - \nu_2 m_2^{*}(s) + \nu_1 V_{11}^{*}(s) + \nu_2 V_{22}^{*}(s) + (\xi_1+\xi_2)V_{12}^{*}(s)$$

$$(4.9)\quad sV_{22}^{*}(s) = (\lambda_2+\mu_2+\nu_2)m_2^{*}(s) + \nu_1 m_1^{*}(s) + 2\xi_2 V_{22}^{*}(s) + 2\nu_1 V_{12}^{*}(s) + \alpha_2 s^{-1}$$

for $\Re(s)$ large enough to ensure the existence of $m_i^{*}(s)$ and $V_{ij}^{*}(s)$ $(i,j=1,2)$. On writing equations (3.4) and (3.5) in the form

$$(4.10)\quad \begin{aligned} sm_1^{*}(s) - a_1 &= \xi_1 m_1^{*}(s) + \nu_2 m_2^{*}(s) + \alpha_1 s^{-1} \\ sm_2^{*}(s) - a_2 &= \xi_2 m_2^{*}(s) + \nu_1 m_1^{*}(s) + \alpha_2 s^{-1} \end{aligned}$$

and combining (4.10) with (4.7) and (4.9) we get

$$(4.11)\quad \begin{aligned} V_{11}^{*}(s)(s-2\xi_1) &= 2\nu_2 V_{12}^{*}(s) + [s+2(\mu_1+\nu_1)]m_1^{*}(s) - a_1 \\ V_{22}^{*}(s)(s-2\xi_2) &= 2\nu_1 V_{12}^{*}(s) + [s+2(\mu_2+\nu_2)]m_2^{*}(s) - a_2 \quad . \end{aligned}$$

We now substitute for $V_{11}^*(s)$ and $V_{22}^*(s)$ from (4.11) into (4.8) to obtain

$$(4.12)\quad V_{12}^*(s)[s-(\xi_1+\xi_2)][(s-2\xi_1)(s-2\xi_2)-4\nu_1\nu_2] = \nu_1[2\lambda_1 m_1^*(s)-a_1](s-2\xi_2) + \nu_2[2\lambda_2 m_2^*(s)-a_2](s-2\xi_1) \quad .$$

But

$$s^2 - 2(\xi_1+\xi_2)s + 4(\xi_1\xi_2-\nu_1\nu_2) = (s-2\omega_1)(s-2\omega_2)$$

whilst

$$\xi_1+\xi_2 = \omega_1+\omega_2 \quad ,$$

and so (4.12) becomes

$$(4.13)\quad V_{12}^*(s) = \nu_1[2\lambda_1 m_1^*(s)-a_1](s-2\xi_2)/(s-2\omega_1)(s-\omega_1-\omega_2)(s-2\omega_2) + \nu_2[2\lambda_2 m_2^*(s)-a_2](s-2\xi_1)/(s-2\omega_1)(s-\omega_1-\omega_2)(s-2\omega_2) \quad .$$

This expression is easily inverted, as from (3.6) and (3.7)

$$(4.14)\quad \begin{aligned} m_1^*(s) &= [a_1 s^2 + (\alpha_1-a_1\xi_2+\nu_2 a_2)s + (\alpha_2\nu_2-\alpha_1\xi_2)]/s(s-\omega_1)(s-\omega_2) \\ m_2^*(s) &= [a_2 s^2 + (\alpha_2-a_2\xi_1+\nu_1 a_1)s + (\alpha_1\nu_1-\alpha_2\xi_1)]/s(s-\omega_1)(s-\omega_2) \end{aligned}$$

and on substituting for $m_1^*(s)$ and $m_2^*(s)$ from (4.14) into expression (4.13), expanding into partial fractions and then using the inverse transformation

$$(s-a)^{-1} \equiv \mathcal{L}[e^{at}] \quad ,$$

we get for $\omega_1 \neq \omega_2$

$$(4.15)\quad V_{12}(t) = r_1 e^{2\omega_1 t} + r_2 e^{(\omega_1+\omega_2)t} + r_3 e^{2\omega_2 t} + r_4 e^{\omega_1 t} + r_5 e^{\omega_2 t} + r_6$$

where r_i $(i=1,\ldots,6)$ are constants listed in Appendix A.

To evaluate $V_{11}(t)$ and $V_{22}(t)$ we write equations (4.3) - (4.5) in the integral form

$$V_{11}(t) = \int_o^t e^{2\xi_1(t-s)} [(\lambda_1+\mu_1+\nu_1)m_1(s)+\nu_2 m_2(s)+2\nu_2 V_{12}(s)+\alpha_1]ds$$

(4.16)
$$V_{22}(t) = \int_o^t e^{2\xi_2(t-s)} [(\lambda_2+\mu_2+\nu_2)m_2(s)+\nu_1 m_1(s)+2\nu_1 V_{12}(s)+\alpha_2]ds$$

and replace $m_1(t)$, $m_2(t)$ and $V_{12}(t)$ by expressions (3.11), (3.12) and (4.15) respectively.

2.4.ii Asymptotic formulae when $\omega_1>0$

If $\omega_1>0$ it follows from (4.15) that for sufficiently large t

(4.17)
$$V_{12}(t) \sim r_1 e^{2\omega_1 t} \qquad (\omega_1 \neq \omega_2)$$

where r_1 is defined in Appendix A. Expressions (4.16) then yield for large t

(4.18)
$$V_{11}(t) \sim [\nu_2 r_1/(\omega_1-\xi_1)]e^{2\omega_1 t} \; ; \; V_{22}(t) \sim [\nu_1 r_1/(\omega_1-\xi_2)]e^{2\omega_1 t}.$$

Thus when t is large the linear correlation coefficient

(4.19)
$$\rho(t) = V_{12}(t)/\{V_{11}(t)V_{22}(t)\}^{\frac{1}{2}} \sim \{(\omega_1^2-\omega_1(\xi_1+\xi_2)+\xi_1\xi_2)/(\nu_1\nu_2)\}^{\frac{1}{2}},$$

i.e. using (3.8)

(4.20)
$$\rho(t) \sim 1 \qquad (\omega_1 > 0;\ \omega_1 \neq \omega_2).$$

2.4.iii Equilibrium formulae when $\omega_1<0$

If $\omega_1<0$ an equilibrium situation develops in which $m_i(t)$ and $V_{ij}(t)$ approach finite non-zero limits as $t\to\infty$. In particular

(3.24)
$$m_i(\infty) = (\alpha_j\nu_j-\alpha_i\xi_j)/(\xi_1\xi_2-\nu_1\nu_2) \qquad (i,j=1,2;\ i\neq j)$$

and, as the $dV_{ij}(t)/dt\to 0$ as $t\to\infty$, equations (4.3)-(4.5) solve to

to give for $\omega_1 \neq \omega_2$

$$(4.21)\quad V_{12}(\infty) = [\alpha_1\nu_1(\lambda_2\nu_2\xi_1 - \lambda_1\xi_2^2) + \alpha_2\nu_2(\lambda_1\nu_1\xi_2 - \lambda_2\xi_1^2)] \div (\xi_1+\xi_2)(\xi_1\xi_2 - \nu_1\nu_2)^2$$

$$(4.22)\quad V_{11}(\infty) = -(\nu_2/\xi_1)V_{12}(\infty) + (\xi_1-\lambda_1)(\alpha_2\nu_2 - \alpha_1\xi_2)/\xi_1(\xi_1\xi_2 - \nu_1\nu_2)$$

$$(4.23)\quad V_{22}(\infty) = -(\nu_1/\xi_2)V_{12}(\infty) + (\xi_2-\lambda_2)(\alpha_1\nu_1 - \alpha_2\xi_1)/\xi_2(\xi_1\xi_2 - \nu_1\nu_2) \quad .$$

The linear correlation coefficient $\rho(\infty)$ does not appear to have a simple form, unlike the previous case $(\omega_1 > 0)$ in which $\rho(\infty) = 1$. Note that (4.21) agrees exactly with the expression for r_6 given in Appendix A.

2.4.iv <u>One-way migration</u>

If migration is allowed to take place in only one direction, so that $\nu_2 = 0$ (say), equations (4.3)-(4.5) may each be integrated in turn to yield

$$V_{11}(t) = \int_o^t e^{2\xi_1(t-s)} [(\lambda_1+\mu_1+\nu_1)m_1(s) + \alpha_1]\,ds$$

$$(4.24)\quad V_{12}(t) = \int_o^t \nu_1 e^{(\xi_1+\xi_2)(t-s)} [V_{11}(s) - m_1(s)]\,ds$$

$$V_{22}(t) = \int_o^t e^{2\xi_2(t-s)} [(\lambda_2+\mu_2)m_2(s) + \nu_1 m_1(s) + 2\nu_1 V_{12}(s) + \alpha_2]\,ds \quad .$$

Now the condition $\nu_2 = 0$ implies

$$\omega_1, \omega_2 = \tfrac{1}{2}[(\xi_1+\xi_2) \pm |\xi_1 - \xi_2|] \quad .$$

Hence, assuming that $\xi_1 > \xi_2$, we have

$$\omega_1 = \xi_1 \quad \text{and} \quad \omega_2 = \xi_2 \quad .$$

If $\xi_1=\xi_2=\xi$ then $\omega_1=\omega_2$ and the solutions to the integrals (4.24) are not too lengthy. In particular if there is no immigration ($\alpha_1=\alpha_2=0$)

$$V_{11}(t) = (\sigma_1 a_1/\xi)e^{\xi t}(e^{\xi t}-1)$$

$$V_{12}(t) = (\nu_1 a_1/\xi)e^{\xi t}(1-e^{\xi t}) + (\nu_1 a_1\sigma_1/\xi^2)e^{\xi t}[1+e^{\xi t}(\xi t-1)] \tag{4.25}$$

$$V_{22}(t) = (\nu_1 a_1+\sigma_2 a_2)\xi^{-1}e^{\xi t}(e^{\xi t}-1) + (a_1\nu_1\sigma_2/\xi^2)e^{\xi t}[e^{\xi t}-(1+\xi t)]$$
$$+ 2(\nu_1^2 a_1/\xi^3)e^{\xi t}\{\xi[e^{\xi t}(1-\xi t)-1] + \sigma_1[e^{\xi t}(\tfrac{1}{2}\xi^2 t^2-\xi t+1)-1]\}$$

where $\sigma_i=\lambda_i+\mu_i+\nu_i$ $(i=1,2)$. Expressions (4.25) were verified by back-substitution into equations (4.3)-(4.5).

The asymptotic behaviour of expressions (4.25) clearly depends on whether $\xi>0$, $\xi=0$ or $\xi<0$. Considering the three cases separately we have the following results.

(i) $\xi>0$:

$$V_{11}(t) \sim (\sigma_1 a_1/\xi)e^{2\xi t}$$

$$V_{12}(t) \sim \nu_1 t(\sigma_1 a_1/\xi)e^{2\xi t}$$

$$V_{22}(t) \sim \nu_1^2 t^2(\sigma_1 a_1/\xi)e^{2\xi t}$$

which gives the asymptotic linear correlation coefficient

$$\rho(t) \sim 1 .$$

(ii) $\xi=0$: equations (4.24) now integrate to give

$$V_{11}(t) = \sigma_1 a_1 t$$

$$V_{12}(t) = -\nu_1 a_1 t + \tfrac{1}{2}\nu_1 a_1\sigma_1 t^2$$

$$V_{22}(t) = (\nu_1 a_1+\sigma_2 a_2)t + a_1\nu_1(\tfrac{1}{2}\sigma_2-\nu_1)t^2 + a_1\nu_1^2\sigma_1 t^3/3$$

which gives the asymptotic linear correlation coefficient

$$\rho(t) \sim \sqrt{(3)}/2 .$$

(iii) $\xi<0$:

$$V_{11}(t) \sim -(\sigma_1 a_1/\xi)e^{\xi t}$$

$$V_{12}(t) \sim (2\lambda_1\nu_1 a_1/\xi^2)e^{\xi t}$$

$$V_{22}(t) \sim -(\nu_1 a_1\sigma_2/\xi)te^{\xi t}$$

which gives the asymptotic linear correlation coefficient

$$\rho(t) \sim (2\lambda_1/\xi)\{\sigma_1\sigma_2 t/\nu_1\}^{-\frac{1}{2}} \quad .$$

2.5 An Alternative Method of Solution for Second-Order Moments

2.5.i General solutions

We have previously seen that as an alternative to solving equations (4.3)-(4.5) by the use of Laplace transforms, we may express the variance-covariance matrix in the integral matrix form (4.6). However, this expression still has to be integrated, and it is preferable to work instead with the equivalent representation for the variance-covariance vector

$$\underline{V}(t) = (V_{11}(t),\ V_{12}(t),\ V_{22}(t))' \ .$$

We may write equations (4.3)-(4.5) in the vector-matrix form

(5.1) $$d\underline{V}(t)/dt = R\underline{V}(t) + \underline{f}(t)$$

where

(5.2) $$R = \begin{bmatrix} 2\xi_1 & 2\nu_2 & 0 \\ \nu_1 & (\xi_1+\xi_2) & \nu_2 \\ 0 & 2\nu_1 & 2\xi_2 \end{bmatrix} \quad ; \quad \underline{f}(t) = \begin{bmatrix} (\lambda_1+\mu_1+\nu_1)m_1(t)+\nu_2 m_2(t)+\alpha_1 \\ -\nu_1 m_1(t)-\nu_2 m_2(t) \\ (\lambda_2+\mu_2+\nu_2)m_2(t)+\nu_1 m_1(t)+\alpha_2 \end{bmatrix},$$

with $\xi_i=\lambda_i-\mu_i-\nu_i$ $(i=1,2)$, and then solve equation (5.1) by using the following lemma (Theorem 1.4 in Bellman (1953), p.14).

LEMMA : If R is a constant matrix, the solution of the equation

$$d\underline{V}(t)/dt = R\underline{V}(t) + \underline{f}(t) \quad , \quad \underline{V}(0) = \underline{c}$$

is given by

$$\underline{V}(t) = \underline{y}(t) + \int_0^t Y(t-s)\underline{f}(s)ds \tag{5.3}$$

where $\underline{y}(t)$ is the solution of the equation

$$d\underline{y}(t)/dt = R\underline{y}(t) \quad , \quad \underline{y}(0) = \underline{c} \tag{5.4}$$

and $Y(t)$ is the solution of the equation

$$dY(t)/dt = RY(t) \quad , \quad Y(0) = I \tag{5.5}$$

the identity matrix.

As $\underline{c} = \underline{V}(0) = \underline{0}$, it follows from (5.4) that

$$\underline{y}(t) \equiv \underline{0} \qquad (t \geq 0) \quad . \tag{5.6}$$

Let us now examine equation (5.5). The eigen-values of the matrix R are given by

$$\zeta_1 = 2\omega_1 \; ; \; \zeta_2 = \omega_1 + \omega_2 \; ; \; \zeta_3 = 2\omega_2$$

where ω_1 and ω_2 are defined in (3.9), and the ζ_i are clearly distinct provided that $\omega_1 \neq \omega_2$. By reference to standard results in matrix theory we see that for $\omega_1 \neq \omega_2$ equation (5.5) integrates to

$$Y(t) = \sum_{i=1}^{3} A_i e^{\zeta_i t} \tag{5.7}$$

(Bartlett (1966), p.27). The matrices A_i are given by

$$A_i = \underline{q}_i \, \underline{r}_i' \qquad (i=1,2,3) \tag{5.8}$$

where $\underline{q}_i$ and $\underline{r}_i'$ are the column and row eigen-vectors of R ,

respectively, normalized so that

$$\underline{r}_i' \underline{q}_i = 1 \qquad (i=1,2,3) \quad . \tag{5.9}$$

But the eigen-vectors of R are

$$\underline{q}_1 = (2\nu_2/(\xi_2-\xi_1+D),\ 1,\ 2\nu_1/(\xi_1-\xi_2+D))'$$

$$\underline{q}_2 = (2\nu_2/(\xi_2-\xi_1),\ 1,\ 2\nu_1/(\xi_1-\xi_2))'$$

$$\underline{q}_3 = (2\nu_2/(\xi_2-\xi_1-D),\ 1,\ 2\nu_1/(\xi_1-\xi_2-D))'$$

$$\underline{r}_1 = (\nu_1/(\xi_2-\xi_1+D),\ 1,\ \nu_2/(\xi_1-\xi_2+D))'$$

$$\underline{r}_2 = (\nu_1/(\xi_2-\xi_1),\ 1,\ \nu_2/(\xi_1-\xi_2))'$$

$$\underline{r}_3 = (\nu_1/(\xi_2-\xi_1-D),\ 1,\ \nu_2/(\xi_1-\xi_2-D))'$$

where

$$D = \{(\xi_1-\xi_2)^2 + 4\nu_1\nu_2\}^{\frac{1}{2}} \quad .$$

Hence on applying the normalization (5.9), we see that expression (5.8) yields

$$A_i = \begin{bmatrix} 2\nu_1\nu_2(\xi_2-\xi_1+jD)^{-2} & 2\nu_2(\xi_2-\xi_1+jD)^{-1} & 2\nu_2^2\{(jD)^2-(\xi_1-\xi_2)^2\}^{-1} \\ \nu_1(\xi_2-\xi_1+jD)^{-1} & 1 & \nu_2(\xi_1-\xi_2+jD)^{-1} \\ 2\nu_1^2\{(jD)^2-(\xi_1-\xi_2)^2\}^{-1} & 2\nu_1(\xi_1-\xi_2+jD)^{-1} & 2\nu_1\nu_2(\xi_1-\xi_2+jD)^{-2} \end{bmatrix} \phi_i \tag{5.10}$$

where $j=2-i$ and

$$\phi_1 = \phi_3 = 2\nu_1\nu_2 D^{-2} \quad ; \quad \phi_2 = (\xi_1-\xi_2)^2 D^{-2} \quad .$$

On combining (5.3), (5.6) and (5.7), we see that

$$\underline{V}(t) = \sum_{i=1}^{3} A_i \int_o^t e^{\zeta_i(t-s)} \underline{f}(s)ds \quad . \tag{5.11}$$

We now write the vector $\underline{f}(s)$ given in expression (5.2) in the form

$$\underline{f}(s) = \underline{l}e^{\omega_1 s} + \underline{m}e^{\omega_2 s} + \underline{n} \quad , \tag{5.12}$$

where $\underline{l}$, $\underline{m}$ and $\underline{n}$ are constant column vectors, and integrate expression (5.11), so obtaining the solution

$$\underline{V}(t) = \sum_{i=1}^{3} A_i[\underline{l}(\omega_1-\zeta_i)^{-1}(e^{\omega_1 t}-e^{\zeta_i t}) + \underline{m}(\omega_2-\zeta_i)^{-1}(e^{\omega_2 t}-e^{\zeta_i t}) + \underline{n}(-\zeta_i)^{-1}(1-e^{\zeta_i t})] \tag{5.13}$$

for $\omega_1 \neq \omega_2$.

2.5.ii Equilibrium formulae when $\omega_1 < 0$

If $\omega_1 < 0$ an equilibrium situation develops in which $m_i(t)$ and $V_{ij}(t)$ approach finite non-zero limits as $t \to \infty$. In particular, it follows from expression (5.13) that

$$\lim_{t\to\infty} \underline{V}(t) = -[\sum_{i=1}^{3} (1/\zeta_i)A_i]\underline{n} = -B\underline{n} \qquad \text{(say)}, \tag{5.14}$$

where we denote the matrix $B \equiv (B_{ij})$ $(i,j=1,2,3)$. But on comparing expressions (3.24), (5.2) and (5.12) we see that the vector $\underline{n}$ may be written in the form

$$\underline{n} = \begin{bmatrix} 2(\xi_1-\lambda_1)(\alpha_1\xi_2-\alpha_2\nu_2) \\ \alpha_1\nu_1(\xi_2-\nu_2) + \alpha_2\nu_2(\xi_1-\nu_1) \\ 2(\xi_2-\lambda_2)(\alpha_2\xi_1-\alpha_1\nu_1) \end{bmatrix} (\omega_1\omega_2)^{-1} \quad . \tag{5.15}$$

Whilst on combining expressions (5.10) and (5.14) we get, after some algebraic manipulation,

$$(B_{21},B_{22},B_{23}) = (-\nu_1\xi_2, 2\xi_1\xi_2, -\nu_2\xi_1)/2\omega_1\omega_2(\xi_1+\xi_2)$$

$$(5.16)\quad (B_{11},B_{12},B_{13}) = -(\nu_2/\xi_1)(B_{21},B_{22},B_{23}) + (1/2\xi_1)(1,0,0)$$

$$(B_{31},B_{32},B_{33}) = -(\nu_1/\xi_2)(B_{21},B_{22},B_{23}) + (1/2\xi_2)(0,0,1) \quad .$$

Substituting for $\underline{n}$ and (B_{ij}) from (5.15) and (5.16) respectively into (5.14) we obtain the solution

$$V_{12}(\infty) = [\alpha_1\nu_1(\lambda_2\nu_2\xi_1 - \lambda_1\xi_2^2) + \alpha_2\nu_2(\lambda_1\nu_1\xi_2 - \lambda_2\xi_1^2)]/(\xi_1+\xi_2)(\xi_1\xi_2-\nu_1\nu_2)^2$$

$$V_{11}(\infty) = -(\nu_2/\xi_1)V_{12}(\infty) + (\xi_1-\lambda_1)(\alpha_2\nu_2-\alpha_1\xi_2)/\xi_1(\xi_1\xi_2-\nu_1\nu_2)$$

$$V_{22}(\infty) = -(\nu_1/\xi_2)V_{12}(\infty) + (\xi_2-\lambda_2)(\alpha_1\nu_1-\alpha_2\xi_1)/\xi_2(\xi_1\xi_2-\nu_1\nu_2)$$

(5.17)

for $\omega_2<\omega_1<0$. This agrees exactly with the previous solution (4.21)-(4.23).

2.5.iii Asymptotic formulae when $\omega_1>0$

If $\omega_1>0$ and $\omega_1>\omega_2$ it follows from (5.13) that the variance-covariance vector

$$(5.18)\qquad \underline{V}(t) \sim e^{2\omega_1 t} A_1[(\omega_1)^{-1}\underline{1} + (2\omega_1-\omega_2)^{-1}\underline{m} + (2\omega_1)^{-1}\underline{n}] \quad .$$

Using the notation of Appendix A, denote

$$d_1 = \alpha_1 - a_1\xi_2 + \nu_2 a_2 \quad ; \quad d_2 = \alpha_2\nu_2 - \alpha_1\xi_2$$

$$f_1 = \alpha_2 - a_2\xi_1 + \nu_1 a_1 \quad ; \quad f_2 = \alpha_1\nu_1 - \alpha_2\xi_1 \quad .$$

Thus we may write expressions (3.11) and (3.12) in the form

$$m_1(t) = (\omega_1-\omega_2)^{-1}\{e^{\omega_1 t}(a_1\omega_1 + d_1 + d_2\omega_1^{-1}) - e^{\omega_2 t}(a_1\omega_2 + d_1 + d_2\omega_2^{-1})\} + d_2(\omega_1\omega_2)^{-1}$$

$$(5.19)$$

$$m_2(t) = (\omega_1-\omega_2)^{-1}\{e^{\omega_1 t}(a_2\omega_1 + f_1 + f_2\omega_1^{-1}) - e^{\omega_2 t}(a_2\omega_2 + f_1 + f_2\omega_2^{-1})\} + f_2(\omega_1\omega_2)^{-1} \quad ,$$

and a comparison of (5.2), (5.12) and (5.19) gives

$$\underline{l}' = \{(2\lambda_1-\xi_1,-\nu_1,\nu_1)(a_1\omega_1+d_1+d_2\omega_1^{-1}) + (\nu_2,-\nu_2,2\lambda_2-\xi_2)(a_2\omega_1+f_1+f_2\omega_1^{-1})\}(\omega_1-\omega_2)^{-1}$$

$$\underline{m}' = -\{(2\lambda_1-\xi_1,-\nu_1,\nu_1)(a_1\omega_2+d_1+d_2\omega_2^{-1}) + (\nu_2,-\nu_2,2\lambda_2-\xi_2)(a_2\omega_2+f_1+f_2\omega_2^{-1})\}(\omega_1-\omega_2)^{-1}$$

$$\underline{n}' = \{(2\lambda_1-\xi_1,-\nu_1,\nu_1)d_2+(\nu_2,-\nu_2,2\lambda_2-\xi_2)f_2\}(\omega_1\omega_2)^{-1} + (\alpha_1,0,\alpha_2) \quad .$$

The matrix A_1 is defined by expression (5.10) with $i=1$, and so expression (5.18) yields

$$V_{12}(t) \sim e^{2\omega_1 t}[\lambda_1\nu_1(\omega_1-\xi_2)(4a_1\omega_1^2+2d_1\omega_1+d_2) + \lambda_2\nu_2(\omega_1-\xi_1)(4a_2\omega_1^2+2f_1\omega_1+f_2)] \div \omega_1^2(2\omega_1-\omega_2)(\omega_1-\omega_2)^2 + Qe^{2\omega_1 t} \tag{5.20}$$

where

$$Q = \{\nu_1[-\xi_1(\omega_1-\xi_2)+\nu_2(\omega_1-\xi_1)-2\nu_1\nu_2](4a_1\omega_1^2+2d_1\omega_1+d_2) + \nu_2[\nu_1(\omega_1-\xi_2)-\xi_2(\omega_1-\xi_1)-2\nu_1\nu_2](4a_2\omega_1^2+2f_1\omega_1+f_2)\}/2\omega_1^2(2\omega_1-\omega_2)(\omega_1-\omega_2)^2 + [\alpha_1\nu_1(\omega_1-\xi_2)+\alpha_2\nu_2(\omega_1-\xi_1)]/2\omega_1(\omega_1-\omega_2)^2 \quad . \tag{5.21}$$

Successive use of the relations

$$\omega_1\omega_2 = \xi_1\xi_2-\nu_1\nu_2 \quad \text{and} \quad \omega_1+\omega_2 = \xi_1+\xi_2$$

reduces Q to the form

$$Q = -[\nu_1a_1(\omega_1-\xi_2) + \nu_2a_2(\omega_1-\xi_1)]/(\omega_1-\omega_2)^2 \quad , \tag{5.22}$$

and the amalgamation of expressions (5.20) and (5.22) agrees exactly with the previous asymptotic solution

$$V_{12}(t) \sim r_1 e^{2\omega_1 t} \qquad (\omega_1 \neq \omega_2) \quad , \tag{4.17}$$

where r_1 is defined in Appendix A.

To determine $V_{11}(t)$ and $V_{22}(t)$ we observe that the rows of the matrix A_1 in expression (5.18) are related by

$$\text{row } 1 = [\nu_2/(\omega_1-\xi_1)] \times \text{row } 2$$

$$\text{row } 3 = [\nu_1/(\omega_1-\xi_2)] \times \text{row } 2 \quad .$$

Hence when t is large

$$V_{11}(t) \sim [\nu_2 r_1/(\omega_1-\xi_1)] e^{2\omega_1 t} \quad ; \quad V_{22}(t) \sim [\nu_1 r_1/(\omega_1-\xi_2)] e^{2\omega_1 t} \tag{5.23}$$

which are the same as expressions (4.18).

2.5.iv Formulae for the homogeneous model

The expressions for the first- and second-order moments simplify considerably for the 'homogeneous' model, in which the parameters are independent of spatial position. Let $\lambda_i=\lambda$, $\mu_i=\mu$, $\nu_i=\nu$ $(i=1,2)$, and for algebraic convenience exclude the immigration component so that $\alpha_1=\alpha_2=0$. Then from (3.11) and (3.12) we have

$$m_i(t) = \tfrac{1}{2}(a_i+a_j)e^{(\lambda-\mu)t} + \tfrac{1}{2}(a_i-a_j)e^{(\lambda-\mu-2\nu)t} \qquad (i=1,2;j\neq i) \tag{5.24}$$

whilst from (5.13) we have

$$\begin{aligned} \underline{V}(t) &= (1/4)(a_1+a_2)(\lambda+\mu)(\lambda-\mu)^{-1}(e^{2(\lambda-\mu)t}-e^{(\lambda-\mu)t})(\underline{e}_1+\underline{e}_2+\underline{e}_3) \\ &\quad + \tfrac{1}{2}(a_1-a_2)(\lambda+\mu)(\lambda-\mu)^{-1}(e^{2(\lambda-\mu-\nu)t}-e^{(\lambda-\mu-2\nu)t})(\underline{e}_1-\underline{e}_3) \\ &\quad + (1/4)(a_1+a_2)(\lambda+\mu+4\nu)(\lambda-\mu-4\nu)^{-1}(e^{2(\lambda-\mu-2\nu)t}-e^{(\lambda-\mu)t})(\underline{e}_1-\underline{e}_2+\underline{e}_3), \end{aligned} \tag{5.25}$$

where the vectors $\underline{e}_i$ are defined by $\underline{e}_1' = (1,0,0)$, $\underline{e}_2' = (0,1,0)$ and $\underline{e}_3' = (0,0,1)$. We see from (5.25) that for $\nu>0$ the

asymptotic linear correlation coefficient is given by $\rho(t) \sim 1$ for $\lambda \geq \mu$; whilst if $\lambda < \mu$ $\rho(t) \sim 4\nu\lambda/(\mu^2+4\nu\mu-\lambda^2)$, even though $V_{11}(t)$, $V_{12}(t)$ and $V_{22}(t)$ all approach zero as $t \to \infty$.

CHAPTER 3

EXACT SOLUTIONS TO THE KOLMOGOROV FORWARD EQUATIONS : TWO-COLONY MODELS

3.1 Introduction

Bartlett (1949) considers the migration of individuals among N colonies in which birth and immigration are absent, and shows that the colony sizes are distributed as a convolution of multinomial distributions. Although this process is conservative (individuals merely move independently of each other from one colony to another), he notes that non-conservatism due to either death or emigration may be covered by the introduction of a further state from which there is no return. He also remarks that immigration can be covered in a similar way by the existence of a large reservoir of individuals capable of passing into the 'proper' colonies of the system. However, he goes on to describe a simpler and more general method of dealing with the problem of immigration. Raman and Chiang (1973) generalize this situation by allowing time-dependent transition rates, but they restrict the matrix of migration rates to be of upper triangular form. They derive a 'solution' in the sense that they write down a system of Riccati equations which give rise to the required probabilities when the system is solved sequentially. Not surprisingly they do not attempt to go further than this.

One application of this process that has produced extensive explicit solutions is the illness-death model treated by Chiang (1968). This model has s states of illness and r states of

death (these states correspond to colonies above). An illness state may be defined to include the absence of illness (health), a single specific disease or stage of disease, or any combination of diseases. A death state is defined by the cause of death, whether single or multiple.

Ruben (1962) allows immigration and obtains the joint factorial moment generating function of the number of individuals in each of the colonies. Wiggins (1960) excludes death and introduces a 'feeding function' $f(t)$ which measures the total number of individuals introduced into the system up to time t. He describes procedures for estimating the migration rates and modifies them to account for the fact that in certain situations the experimenter observes concentrations of individuals in a compartment (colony) rather than total numbers. Often, in practice, individual compartments are inaccessible for observation and instead time series data are available on the passage of material to the system exterior only. Techniques for estimating the migration and death parameters in these circumstances are described in Matis and Hartley (1971).

When births are allowed, this process becomes far more difficult to analyse. Puri (1968) solves the stochastic equations in two special cases and determines a sequence of stochastic processes whose generating functions converge to the required probability generating function. In Chapter Two I derived the Kolmogorov forward differential equation for the p.g.f. $G(z_1,z_2;t)$, and used it to determine the first- and second-order moments of the process. The equation cannot be solved directly and the analysis

of $G(z_1,z_2;t)$ in Chapter Five involves the use of two different iterative techniques on the corresponding backward equations. First I obtain a power series solution in terms of the common migration rate ν. Then I show how to generate sequences of functions, defined recursively by means of integral equations, that converge to the required p.g.f.

Exact expressions for $G(z_1,z_2;t)$ may be obtained for certain special cases in which some of the parameters are equated to zero. In this chapter I obtain exact solutions for three such cases, and develop a recursive solution for the probabilities $p_{ij}(t)$ in a fourth.

3.2 Model 1 $(\lambda_1 = \lambda_2 = 0)$

3.2.i Lagrange's linear equation

In order to obtain exact solutions to the Kolmogorov forward equation, namely

$$\begin{aligned}\frac{\partial G}{\partial t} &= [(\lambda_1 z_1 - \mu_1)(z_1-1)+\nu_1(z_2-z_1)]\frac{\partial G}{\partial z_1} \\ &\quad +[(\lambda_2 z_2-\mu_2)(z_2-1)+\nu_2(z_1-z_2)]\frac{\partial G}{\partial z_2}+[\alpha_1(z_1-1)+\alpha_2(z_2-1)]G \quad ,\end{aligned} \tag{2.2.2}$$

we shall treat it as a Lagrange equation. The method of solution of this type of differential equation is described in detail in Piaggio (1962, Chapter 12), and so I shall present only a brief outline of the technique used and refer the reader to Piaggio's text for the full discussion of the analytical implications.

In its simplest form the Lagrange equation may be written as

$$P\frac{\partial z}{\partial x} + Q\frac{\partial z}{\partial y} = R \tag{2.1}$$

where P,Q,R are functions of x,y,z . The simultaneous equations

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R} \tag{2.2}$$

are called the auxiliary equations and represent a family of curves such that the tangent at any point has direction-cosines in the ratio $P:Q:R$. Let

$$u = \text{const.} \quad \text{and} \quad v = \text{const.}$$

be any two independent solutions of these simultaneous equations. Then

$$f(u,v) = 0$$

is an integral of (2.1) if f is any arbitrary function. This is called the General Integral of Lagrange's Linear Equation. The form of the function f can be determined provided that the initial conditions are known. Sufficient conditions for the validity of the above reasoning are that all the first-order partial derivatives (such as $\partial u/\partial x$, etc.,) are continuous, and that P,Q,R do not all vanish simultaneously. The linear equation with n independent variables $(n=2,3,\ldots)$ is solved in an analogous manner. The Kolmogorov equation (2.2.2) considered in this chapter, for example, has three independent variables.

3.2.ii Probability generating function

The auxiliary equations corresponding to the forward differential equation (2.2.2) are as follows

$$\frac{dt}{1}=\frac{-dz_1}{[(\lambda_1 z_1-\mu_1)(z_1-1)+\nu_1(z_2-z_1)]}=\frac{-dz_2}{[(\lambda_2 z_2-\mu_2)(z_2-1)+\nu_2(z_1-z_2)]}$$

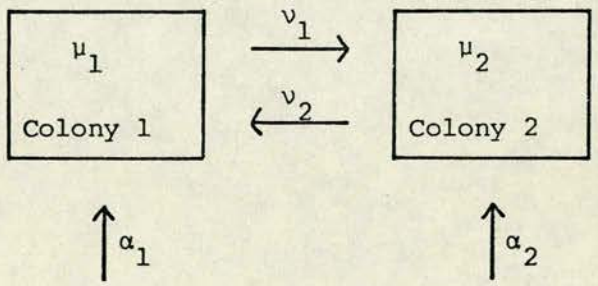
(2.3)

$$=\frac{dG}{G[\alpha_1(z_1-1)+\alpha_2(z_2-1)]} \quad .$$

Although these equations are intractable as they stand, they may be rendered soluble when suitable parameters are equated to zero. As our first example we shall exclude birth from the process by putting $\lambda_1=\lambda_2=0$ (this model is illustrated in figure 3.1). All three equations are now linear in z_1 and z_2, and hence may be integrated in a straightforward manner.

Figure 3.1

Two-colony process with zero birth rate



Denote $\sigma_i = \mu_i+\nu_i$ $(i=1,2)$, and write

$$y_i = 1-z_i \qquad (i=1,2) \quad .$$

Combining the first two equations in (2.3) we have, for any constant b,

$$(2.4) \qquad dt = \frac{d(y_1-by_2)}{y_1(\sigma_1+\nu_2 b)-y_2(\nu_1+\sigma_2 b)}$$

$$= -\frac{d(y_1-by_2)}{(\sigma_1+\nu_2 b)[\,y_1-y_2(\nu_1+\sigma_2 b)/(\sigma_1+\nu_2 b)]} \quad .$$

In order to make the numerator the differential of the denominator consider the equation

$$b = (\nu_1+\sigma_2 b)/(\sigma_1+\nu_2 b)$$

i.e.
$$\nu_2 b^2 + (\sigma_1-\sigma_2)b - \nu_1 = 0 \quad .$$

This has roots

$$(2.5) \qquad b_1,b_2 = (1/2\nu_2)[\,(\sigma_2-\sigma_1) \pm \{(\sigma_1-\sigma_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}}] \quad .$$

Hence if $b=b_1$ or $b=b_2$, equation (2.4) integrates directly to give

$$(2.6) \qquad \begin{aligned} y_1 - b_1 y_2 &= B_1 e^{r_1 t} \\ y_1 - b_2 y_2 &= B_2 e^{r_2 t} \end{aligned}$$

where

$$(2.7) \qquad r_1,r_2 = \tfrac{1}{2}[\,(\sigma_1+\sigma_2) \pm \{(\sigma_1-\sigma_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}}]$$

and B_1,B_2 are constants of integration. Note that if $b_1=b_2$ then $r_1=r_2$ and expressions (2.6) are equivalent.

Returning to the auxiliary equations (2.3) we have

$$(2.8) \qquad dG/G = -(\alpha_1 y_1+\alpha_2 y_2)dt \quad .$$

Solving equations (2.6) for y_1 and y_2 we get, for $b_1 \neq b_2$,

$$
\begin{aligned}
y_1 &= (b_2-b_1)^{-1}(b_2B_1e^{r_1t} - b_1B_2e^{r_2t}) \\
y_2 &= (b_2-b_1)^{-1}(B_1e^{r_1t} - B_2e^{r_2t}) \quad ,
\end{aligned}
\tag{2.9}
$$

and on substituting for y_1 and y_2 from (2.9) into (2.8) and integrating the resulting expression we have

$$
\begin{aligned}
G(z_1,z_2;t) = \text{const.}\exp\{(b_1-b_2)^{-1}[r_1^{-1}(\alpha_1b_2+\alpha_2)B_1e^{r_1t} \\
- r_2^{-1}(\alpha_1b_1+\alpha_2)B_2e^{r_2t}]\} \quad .
\end{aligned}
\tag{2.10}
$$

When $b_1 \neq b_2$ the two solutions (2.6) are different and may be combined with (2.10) to give the general solution

$$
\begin{aligned}
G(z_1,z_2;t) = f(B_1,B_2)\exp\{(b_1-b_2)^{-1}[r_1^{-1}(\alpha_1b_2+\alpha_2)B_1e^{r_1t} \\
- r_2^{-1}(\alpha_1b_1+\alpha_2)B_2e^{r_2t}]\}
\end{aligned}
\tag{2.11}
$$

where f is an arbitrary function.

Expression (2.11) is the general integral of Lagrange's linear equation, and to derive the required integral we find that solution (2.11) which satisfies the initial condition

$$
G(z_1,z_2;0) = z_1^{a_1}z_2^{a_2} \quad . \tag{2.12}
$$

At $t=0$ expressions (2.9) become

$$
\begin{aligned}
z_1 &= 1 - (b_2-b_1)^{-1}(b_2B_1-b_1B_2) \\
z_2 &= 1 - (b_2-b_1)^{-1}(B_1-B_2) \quad ,
\end{aligned}
$$

and so

$$f(B_1,B_2) = [\,1-(b_2-b_1)^{-1}(b_2B_1-b_1B_2)]^{a_1}[\,1-(b_2-b_1)^{-1}(B_1-B_2)]^{a_2}$$

$$\times \exp\{-(b_1-b_2)^{-1}[\,r_1^{-1}(\alpha_1b_2+\alpha_2)B_1 - r_2^{-1}(\alpha_1b_1+\alpha_2)B_2]\} \quad .$$

We now replace the constants B_1 and B_2 by expressions (2.6), namely

$$B_1 = (y_1-b_1y_2)e^{-r_1t} \quad ; \quad B_2 = (y_1-b_2y_2)e^{-r_2t} \quad ,$$

and obtain the following solution $(b_1 \neq b_2)$ for $G(z_1,z_2;t)$

$$G(z_1,z_2;t) = [\,1-(b_2-b_1)^{-1}\{(1-z_1)(b_2e^{-r_1t}-b_1e^{-r_2t})$$

$$(2.13) \qquad - (1-z_2)b_1b_2(e^{-r_1t}-e^{-r_2t})\}]^{a_1}$$

$$\times [\,1-(b_2-b_1)^{-1}\{(1-z_1)(e^{-r_1t}-e^{-r_2t})-(1-z_2)(b_1e^{-r_1t}-b_2e^{-r_2t})\}]^{a_2}$$

$$\times \exp\{(b_2-b_1)^{-1}[\,r_1^{-1}(\alpha_1b_2+\alpha_2)\{(1-z_1)-b_1(1-z_2)\}(e^{-r_1t}-1)$$

$$- r_2^{-1}(\alpha_1b_1+\alpha_2)\{(1-z_1)-b_2(1-z_2)\}(e^{-r_2t}-1)]\} \quad .$$

Expression (2.13) was verified by back-substitution into equation (2.2.2). Note that $G(1,1;t) \equiv 1$ $(t \geq 0)$ as required.

3.2.iii Equilibrium distribution

The equilibrium distribution is easily derived by letting $t\to\infty$ in (2.13). For, writing (2.7) in the form

$$r_1,r_2 = \tfrac{1}{2}[\,(\sigma_1+\sigma_2) \pm \{(\sigma_1+\sigma_2)^2-4[\,(\mu_1+\nu_1)(\mu_2+\nu_2)-\nu_1\nu_2]\}^{\frac{1}{2}}] \quad ,$$

we see that $r_1 > 0$ and $r_2 > 0$ (except for the trivial case $\mu_1=\mu_2=0$) . Substituting for b_1 and b_2 from (2.5) into (2.13)

and letting $t\to\infty$ we get the equilibrium p.g.f.

$$(2.14)\quad G(z_1,z_2;\infty) = \exp\{-(\sigma_1\sigma_2-\nu_1\nu_2)^{-1}[(1-z_1)(\alpha_2\nu_2+\alpha_1\sigma_2) + (1-z_2)(\alpha_1\nu_1+\alpha_2\sigma_1)]\} \quad .$$

Expression (2.14) was verified by back-substitution into equation (2.2.2).

It follows from (2.14) that in equilibrium the colony sizes are distributed as a bivariate Poisson variable with parameters

$$(2.15)\quad (\sigma_1\sigma_2-\nu_1\nu_2)^{-1}(\alpha_2\nu_2+\alpha_1\sigma_2) \quad \text{and} \quad (\sigma_1\sigma_2-\nu_1\nu_2)^{-1}(\alpha_1\nu_1+\alpha_2\sigma_1) \quad ,$$

respectively. Now from Cox and Miller (1965, p.168) we see that for the single-colony simple immigration-death process with immigration and death parameters α and μ respectively, the equilibrium p.g.f. of population size is given by

$$(2.16)\quad G(z;\infty) = \exp\{(\alpha/\mu)(z-1)\} \quad .$$

A comparison of (2.14) and (2.16) then shows that in equilibrium the colonies behave as if they were isolated from each other, the ratios of their immigration rates to death rates being as $(\alpha_2\nu_2+\alpha_1\sigma_2)$ to $(\sigma_1\sigma_2-\nu_1\nu_2)$ and $(\alpha_1\nu_1+\alpha_2\sigma_1)$ to $(\sigma_1\sigma_2-\nu_1\nu_2)$ for colonies 1 and 2 respectively.

3.2.iv Analogous processes

Bartlett considers a similar situation in which birth, death and immigration are suppressed but individuals may migrate independently between N colonies instead of just 2. He derives the equivalent solution to (2.13) in the vector-matrix form

$$G(\underline{z};t) = [\underline{\eta}_1'\underline{z} + \sum_{s=2}^{N} c_s \underline{\eta}_s'\underline{z} e^{\phi_s t}]^n \qquad (2.17)$$

(Bartlett (1949), expression (39)). Here ϕ_s (s=1,...,N) are the eigen-values of the transition matrix

$$\begin{bmatrix} -\Sigma_s \nu_{1s} & \nu_{12} & \nu_{13} & \cdots \\ \nu_{21} & -\Sigma_s \nu_{2s} & \nu_{23} & \cdots \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{bmatrix} ,$$

$\underline{\eta}_s$ are the corresponding column eigen-vectors, $\underline{z}' = (z_1,\ldots,z_N)$, $G(\underline{z};0) = z_1^n$, and the c_s are defined by $z_1 = \Sigma_s c_s \underline{\eta}_s'\underline{z}$. When N=2 expression (2.17) may be considerably simplified. For on placing $\alpha_1=\alpha_2=0$, $\mu_1=\mu_2=0$, $a_1=n$ and $a_2=0$ in expression (2.13) we get

$$G(z_1,z_2;t) = [1 - (\nu_1+\nu_2)^{-1}\{(1-z_1)(\nu_2+\nu_1 e^{-t}) + \nu_1(1-z_2)(1-e^{-t})\}]^n . \qquad (2.18)$$

Note that in this particular case we obtain the Ehrenfest model of diffusion (Cox and Miller (1965), p.130-132) by observing the sizes of the two colonies after each migration.

If we wish to follow the progress of one particular individual that is situated in colony 1 (say) at time t=0, we put $\alpha_1=\alpha_2=0$, $a_1=1$ and $a_2=0$ in expression (2.13). This yields the probabilities

$$p_{10}(t) = (b_2-b_1)^{-1}(b_2 e^{-r_1 t} - b_1 e^{-r_2 t})$$

$$p_{01}(t) = (b_1-b_2)^{-1} b_1 b_2 (e^{-r_1 t} - e^{-r_2 t}) \qquad (2.19)$$

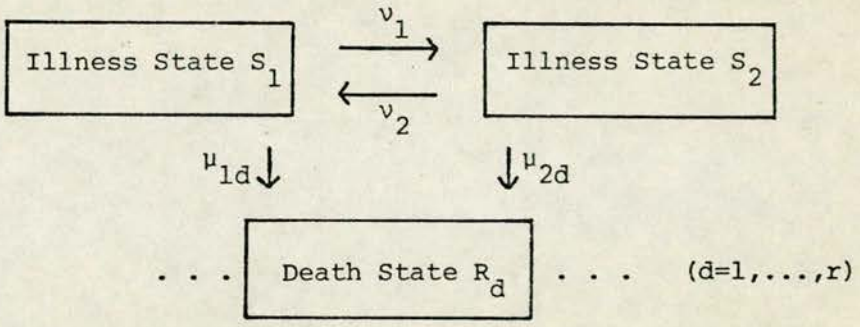
$$p_{00}(t) = 1 + (b_1-b_2)^{-1}[e^{-r_1 t}(b_2-b_1b_2) + e^{-r_2 t}(-b_1+b_1b_2)] \quad .$$

Chiang (1968) suppresses immigration but his model, which is illustrated in figure 3.2, has r different states of death R_d $(d=1,\ldots,r)$ and 2 different states of illness S_i $(i=1,2)$ with

Pr[an individual in the illness state S_i at time t 'migrates' to the death state R_d in $(t,t+\delta t)$] $= \mu_{id}\delta t + o(\delta t)$ $(i=1,2;d=1,\ldots,r)$.

Figure 3.2

Chiang's illness-death process with 2 illness states and r death states



Suppose that the population at time $t=0$ consists of a single individual in state S_1, and let $p_{10}(t)$, $p_{01}(t)$ and $p_{00}(t)$ denote the probabilities that this individual is in either state S_1, S_2 or R_d, respectively, at time $t \geq 0$. He derived the solution

$$p_{10}(t) = \sum_{i=1}^{2} \frac{(\rho_i+\sigma_2)}{(\rho_i-\rho_j)} e^{\rho_i t}$$

$$(2.20) \qquad p_{01}(t) = \sum_{i=1}^{2} \frac{\nu_1}{(\rho_i-\rho_j)} e^{\rho_i t}$$

$$p_{00}(t) = \sum_{i=1}^{2} \frac{1}{\rho_i(\rho_i-\rho_j)} (e^{\rho_i t}-1)[(\rho_i+\sigma_2)\mu_{1d}+\nu_1\mu_{2d}]$$

(Chiang (1968), Chapter 4, expressions (2.25) and (2.30)) by assuming that the required probabilities are linear mixtures of exponentials. Here $j\neq i$, $d=1,\ldots,r$ and $\rho_i=-r_j$. A little algebra shows that expressions (2.19) and (2.20) are equivalent when $d=1$. Chiang also derived the expected time of duration in illness and death and he studied the number of transitions made in a given time interval.

3.2.v One-way migration

From expression (2.11) onwards the theory is developed under the assumption that $b_1\neq b_2$. However, it follows from (2.5) that this condition is equivalent to the assumption that

$$\sigma_1 = \sigma_2 \quad \text{and} \quad \nu_1\nu_2 = 0$$

do not hold simultaneously. Let us now assume that they do and put

$$\nu_2 = 0$$

together with

$$\sigma_1 = \sigma_2$$

i.e.
$$\mu_1 + \nu_1 = \mu_2 \ .$$

The first two of equations (2.3) integrate to give

$$(z_2-1)e^{-\mu_2 t} = \text{const.}$$

$$[(z_1-1) + \nu_1 t(z_2-1)]e^{-\mu_2 t} = \text{const.} \ ,$$

whilst the remaining analysis carries through in the same manner as before. We find that for $b_1=b_2$ the solution for $G(z_1,z_2;t)$ is given by the expression

$$G(z_1,z_2;t) = [1-(1-z_1)e^{-\mu_2 t} - \nu_1 t(1-z_2)e^{-\mu_2 t}]^{a_1} [1-(1-z_2)e^{-\mu_2 t}]^{a_2}$$

$$\times \exp\{\mu_2^{-1}\alpha_1(1-e^{-\mu_2 t})(z_1-1)\}$$

$$\times \exp\{\mu_2^{-1}[(\alpha_2+\alpha_1\nu_1\mu_2^{-1})(1-e^{-\mu_2 t}) - \alpha_1\nu_1 te^{-\mu_2 t}](z_2-1)\} . \quad (2.21)$$

Expression (2.21) was verified by back-substitution into equation (2.2.2). Note that $G(1,1;t) \equiv 1$ for all $t \geq 0$ as required.

Raman and Chiang (1973) subsequently generalized this situation by allowing time-dependent transition rates, with migration occurring between a finite number of colonies denoted by $i=1,\ldots,N$. They consider one-way migration in the sense that a migration may take place from colony i to colony j only if $i < j$ $(i,j=1,\ldots,N)$. The authors also discuss an application of their results to a problem arising in a study of the epidemiology of leprosy.

The equilibrium p.g.f.

$$G(z_1,z_2;\infty)=\exp\{(\alpha_1/(\mu_1+\nu_1))(z_1-1)\}\exp\{(1/\mu_2)[\alpha_2+\alpha_1\nu_1/(\mu_1+\nu_1)](z_2-1)\} \quad (2.22)$$

may be obtained either by letting $t\to\infty$ in (2.21) or by putting $\sigma_1=\sigma_2$ and $\nu_2=0$ in (2.14).

If we compare expressions (2.16) and (2.22) we see that in equilibrium colony 1 develops as though it were an isolated colony with an effective death rate $(\mu_1+\nu_1)$, as would be expected, whilst

colony 2 develops as though it were an isolated colony with an effective immigration rate $\alpha_2+\alpha_1\nu_1/(\mu_1+\nu_1)$. This latter result may be interpreted as follows. Immigrants to colony 1 either die at rate μ_1 or migrate to colony 2 at rate ν_1 , and so the proportion migrating is equal to $\nu_1/(\mu_1+\nu_1)$. Individuals therefore enter colony 2 via colony 1 at rate $\alpha_1\nu_1/(\mu_1+\nu_1)$, giving a total rate of immigration to colony 2 of $\alpha_2+\alpha_1\nu_1/(\mu_1+\nu_1)$.

3.3 Model 2 $(\lambda_1 = \nu_2 = 0)$

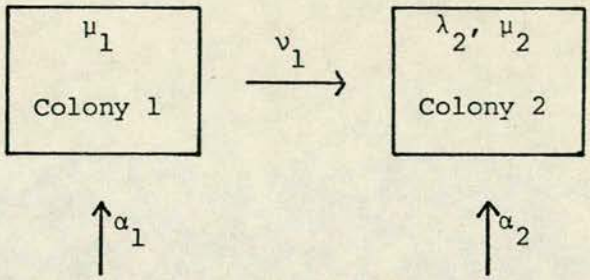
3.3.i Probability generating function

If $\lambda_1 > 0$ and $\lambda_2 > 0$ the first two of the auxiliary equations (2.3) are each quadratic in z_1 and z_2 , and in general this renders the forward differential equation (2.2.2) difficult to solve. Exceptions include, for example, models with $\nu_1=\nu_2=0$, in which case the colonies are unconnected, and $\mu_1=\mu_2=\alpha_1=\alpha_2=\nu_2=0$ which is discussed in section 3.5.

As my second model I shall assume that $\lambda_2 > 0$ together with the added restriction that $\lambda_1=\nu_2=0$. Colony 1 now develops as a simple immigration-death process with parameters α_1 and $(\mu_1+\nu_1)$ respectively. Colony 2 develops as an immigration-birth-death process in which the immigration component consists of direct immigration at rate α_2 together with indirect entry via colony 1. This situation is illustrated in figure 3.3.

Figure 3.3

Two-colony process with $\lambda_1 = \nu_2 = 0$



Under the above restrictions the first two of the auxiliary equations (2.3) become

$$dz_1/dt = -\mu_1(1-z_1) - \nu_1(z_2-z_1) \tag{3.1}$$

$$dz_2/dt = -(\lambda_2 z_2-\mu_2)(z_2-1) \quad . \tag{3.2}$$

Equation (3.2) integrates directly to give

$$[(z_2-1)/(\lambda_2 z_2-\mu_2)]e^{(\lambda_2-\mu_2)t} = A \quad , \tag{3.3}$$

where A is a constant of integration, and this expression may be rewritten as

$$(z_2-1) = A(\lambda_2-\mu_2)/[e^{(\lambda_2-\mu_2)t} - A\lambda_2] \quad . \tag{3.4}$$

Substituting for (z_2-1) from (3.4) into (3.1) we get

$$-dz_1/dt = (\mu_1+\nu_1)(1-z_1) + \nu_1 A(\lambda_2-\mu_2)[e^{(\lambda_2-\mu_2)t} - A\lambda_2]^{-1} \quad ,$$

whence

$$\frac{d}{dt}[(1-z_1)e^{(\mu_1+\nu_1)t}] = \nu_1 A(\lambda_2-\mu_2)e^{(\mu_1+\nu_1)t}[e^{(\lambda_2-\mu_2)t} - A\lambda_2]^{-1} \quad . \tag{3.5}$$

For convenience write

$$q = \mu_2 - \lambda_2 \quad \text{and} \quad r = \mu_1 + \nu_1 = \sigma_1 \ .$$

Then equation (3.5) integrates to give

$$(3.6) \qquad (1-z_1)e^{-rt} = B - \int^t \frac{\nu_1 Aqe^{(q-r)s}\,ds}{1-\lambda_2 Ae^{qs}}$$

where B is a constant of integration.

In view of results (3.4) and (3.6), the third auxiliary equation in (2.3) may be written as

$$\frac{d}{dt}(\log G) = -\alpha_1 Be^{rt} + \alpha_1 e^{rt} \int^t \frac{\nu_1 Aqe^{(q-r)s}\,ds}{1-\lambda_2 Ae^{qs}} - \frac{\alpha_2 Aqe^{qt}}{1-A\lambda_2 e^{qt}}$$

which integrates to give

$$(3.7) \quad G = \text{const.}\,(1-A\lambda_2 e^{qt})^{\alpha_2/\lambda_2} \exp\{-(\alpha_1/r)Be^{rt} + \alpha_1 \int^t e^{rs} \int^s \frac{\nu_1 qAe^{(q-r)\tau}\,d\tau\,ds}{1-\lambda_2 Ae^{q\tau}}\} \ .$$

Thus the general solution for $G(z_1, z_2; t)$ may be written in the form

$$(3.8) \quad G(z_1, z_2; t) = f(A,B)\ (1-A\lambda_2 e^{qt})^{\alpha_2/\lambda_2} \exp\{-(\alpha_1/r)Be^{rt} + \alpha_1 \int^t e^{rs} \int^s \frac{\nu_1 qAe^{(q-r)\tau}\,d\tau\,ds}{1-\lambda_2 Ae^{q\tau}}\}$$

where f is an arbitrary function.

For algebraic simplicity let us assume that both colonies are initially empty. Then at $t=0$

$$G(z_1, z_2; 0) = 1$$

and so (3.8) becomes

$$1 = f(A,B)\,(1-A\lambda_2)^{\alpha_2/\lambda_2} \exp\{-(\alpha_1/r)B + \alpha_1 \int_o^{\cdot} e^{rs} \int^s \frac{\nu_1 qAe^{(q-r)\tau}}{1-\lambda_2 Ae^{q\tau}} d\tau ds\}$$

whence

$$(3.9)\quad G(z_1,z_2;t) = \left[\frac{1-A\lambda_2 e^{qt}}{1-A\lambda_2}\right]^{\alpha_2/\lambda_2} \exp\{-(\alpha_1/r)B(e^{rt}-1) + \alpha_1 \int_o^t e^{rs} \int^s \frac{\nu_1 qAe^{(q-r)\tau}}{1-\lambda_2 Ae^{q\tau}} d\tau ds\}\,.$$

Before we replace the constants A and B in expression (3.9) let us define the indefinite integral

$$(3.10)\qquad J(t) = \int^t \frac{P(z_2)e^{(r-q)s}}{1-\lambda_2 P(z_2)e^{-qs}} ds$$

where

$$(3.11)\qquad P(z_2) = (z_2-1)/(\lambda_2 z_2 - \mu_2)\,.$$

[Note that for typographical reasons I sometimes refrain from quoting functions in full. For example, $J(t)$ in expression (3.10) is a function of both z_2 and t].

Then expression (3.3) is equivalent to

$$(3.12)\qquad P(z_2)e^{(\lambda_2-\mu_2)t} = A\,,$$

and on substituting for A from (3.12) we see that expression (3.6) is equivalent to

$$(3.13)\qquad Q(z_1,z_2)e^{-rt} = B$$

where

$$(3.14) \qquad Q(z_1,z_2) = (1-z_1) + e^{rt}\int^t \frac{\nu_1 q e^{(q-r)s} P e^{-qt}\, ds}{1-\lambda_2 e^{qs} P e^{-qt}} \quad .$$

On putting $\tau=t-s$ we may write

$$Q(z_1,z_2) = (1-z_1) - \nu_1 q \int^0 \frac{P e^{(r-q)\tau}\, d\tau}{1-\lambda_2 P e^{-q\tau}}$$

which, when we use (3.10), gives

$$(3.15) \qquad Q(z_1,z_2) = (1-z_1) - \nu_1 q J(0) \quad .$$

Finally, let us consider the integral in expression (3.9). Denote

$$I(t) = \int_o^t e^{rs}\int^s \frac{A e^{(q-r)\tau}\, d\tau ds}{1-\lambda_2 A e^{q\tau}} = e^{-qt}\int_o^t e^{rs}\int^s \frac{P e^{(q-r)\tau}\, d\tau ds}{1-\lambda_2 P e^{q(\tau-t)}} \quad .$$

If we put $u=t-\tau$ and $v=t-s$ then this integral simplifies to give

$$(3.16) \qquad I(t) = -\int_o^t J(v) e^{-rv}\, dv \quad ,$$

and integrating (3.16) by parts yields

$$(3.17) \quad I(t) = (1/r)(e^{-rt}J(t)-J(0)) + (1/rq\lambda_2)[\log(1-\lambda_2 P) - \log(1-\lambda_2 P e^{-qt})] \quad .$$

Denote the indefinite integral

$$(3.18) \qquad J_1(t) \equiv J(t) - J(0) \quad .$$

Then on substituting for A and B in (3.9) from expressions (3.10)-(3.17) we obtain, after a little algebra, the solution

$$(3.19)\quad G(z_1,z_2;t) = \left[\frac{1-\lambda_2 P}{1-\lambda_2 P e^{-qt}}\right]^{\left[\frac{\alpha_2}{\lambda_2}+\frac{\alpha_1\nu_1}{\lambda_2 r}\right]} \exp\{(\alpha_1\nu_1 q/r)e^{-rt}J_1(t)\} \times \exp\{(\alpha_1/r)(z_1-1)(1-e^{-rt})\}\ .$$

Expression (3.19) was verified by back-substitution into equation (2.2.2). Note that $G(1,1;t) \equiv 1$ for all $t \geq 0$ as required. Puri (1968) obtains the equivalent expression to (3.19) for the case in which immigration is excluded but all the remaining rates are time-dependent.

The p.g.f. $G(z_1,z_2;t)$ clearly factorizes into the product

$$G(z_1,z_2;t) = G_1(z_1;t)G_2(z_2;t)$$

where $G_1(z_1;t)$ and $G_2(z_2;t)$ are independent of z_2 and z_1 respectively. In particular,

$$G_1(z_1;t) = \exp\{(\alpha_1/(\mu_1+\nu_1))(z_1-1)(1-e^{-(\mu_1+\nu_1)t})\}$$

and this is the p.g.f. of a single colony undergoing a simple immigration-death process with parameters α_1 and $(\mu_1+\nu_1)$ respectively (Cox and Miller (1965), p.168). Now the first part of $G_2(z_2;t)$ is

$$G_2^{(1)}(z_2;t) = [(1-\lambda_2 P)/(1-\lambda_2 P e^{-qt})]^{\alpha_2/\lambda_2}\ .$$

Hence on using (3.11) we have

$$G_2^{(1)}(z_2;t) = \left[\frac{\lambda_2-\mu_2}{\lambda_2 e^{-(\mu_2-\lambda_2)t}-\mu_2}\right]^{\alpha_2/\lambda_2}\left[1-\frac{\lambda_2 z_2(e^{-(\mu_2-\lambda_2)t}-1)}{\lambda_2 e^{-(\mu_2-\lambda_2)t}-\mu_2}\right]^{-\alpha_2/\lambda_2}$$

which is the p.g.f. of a simple immigration-birth-death process with parameters α_2, λ_2 and μ_2 respectively (Bharucha-Reid (1960), p.174 - note the mistake in his result (4.34) in which a pair of brackets are omitted). But it follows from (3.19) that

$$G(z_1,z_2;t) \equiv G_1(z_1;t)G_2^{(1)}(z_2;t)G_2^{(2)}(z_2;t) \tag{3.20}$$

where

$$G_2^{(2)}(z_2;t) = [(1-\lambda_2 P)/(1-\lambda_2 P e^{-qt})]^{(\alpha_1\nu_1/\lambda_2 r)} \times \exp\{(\alpha_1\nu_1 q/r)e^{-rt}J_1(t)\} \quad . \tag{3.21}$$

If migration is not present $(\nu_1=0)$

$$G_2^{(2)}(z_2;t) \equiv 1 \qquad (t \geq 0) \quad .$$

Hence the effect of introducing migration at rate ν_1 from colony 1 to colony 2 is to multiply the joint p.g.f.

$$G_1(z_1;t)G_2^{(1)}(z_2;t)$$

by the generating function (3.21), and to increase the 'death' rate in colony 1 from μ_1 to $\mu_1+\nu_1$.

3.3.ii Special case $(\lambda_1=\nu_2=0,\ \lambda_2=\mu_2)$

When $\lambda_2=\mu_2$ the solution (3.19) may be replaced by a simpler expression. For writing

$$K = \left[\frac{\alpha_2}{\lambda_2} + \frac{\alpha_1\nu_1}{\lambda_2 r}\right] \quad ,$$

and remembering that $q=\mu_2-\lambda_2$, we have

$$\lim_{q\to 0}[(1-\lambda_2 P)/(1-\lambda_2 P e^{-qt})]^K = [1+\lambda_2 t(1-z_2)]^{-K} \tag{3.22}$$

whilst

$$(3.23)\qquad \lim_{q\to 0}[qJ_1(t)] = \int_0^t \frac{(1-z_2)e^{rs}\,ds}{1+\lambda_2 s(1-z_2)} \equiv J_2(t) \qquad \text{(say)} \quad .$$

Comparing (3.19), (3.22) and (3.23) we obtain the solution $(\lambda_2=\mu_2)$

$$(3.24)\qquad G(z_1,z_2;t) = [1+\lambda_2 t(1-z_2)]^{-K}\exp\{(\alpha_1/r)(z_1-1)(1-e^{-rt})\}$$
$$\times \exp\{(\alpha_1\nu_1/r)e^{-rt}J_2(t)\} \quad .$$

Expression (3.24) was verified by back-substitution into the differential equation (2.2.2).

3.3.iii Solution in closed form

If $r=nq$ for some integer n, the integral $J_1(t)$ in the solution (3.19) may be replaced by an expression in closed form. We may take $n\neq 0$, for $n=0$ implies $r=\mu_1+\nu_1=0$ whence $\mu_1=\nu_1=0$ and the problem is trivial.

Substituting for

$$x = \lambda_2 P e^{-qs}$$

in (3.18) with $r=nq$ we have

$$J_1(t) = -(1/\lambda_2 q)(\lambda_2 P)^n \int_{s=0}^{s=t} x^{-n}(1-x)^{-1}dx \quad ,$$

and this expression integrates to give

$$(3.25)\qquad \exp\{(\alpha_1\nu_1 q/r)e^{-rt}J_1(t)\} = \left[\frac{1-\lambda_2 P}{1-\lambda_2 P e^{-qt}}\right]^{-c} E(z_2;t)$$

where

$$E(z_2;t) = \begin{cases} \exp\{c[\,qt + \sum_{i=1}^{n-1} (\lambda_2 P)^{-i} i^{-1}(e^{iqt}-1)]\,\} & (q=\mu_2-\lambda_2>0) \\ \exp\{c \sum_{i=-n}^{1} (\lambda_2 P)^{i} i^{-1}(e^{-iqt}-1)\} & (q=\mu_2-\lambda_2<0) \end{cases} .$$

Here

$$c = (\alpha_1\nu_1/\lambda_2 r)(\lambda_2 P)^n e^{-rt} ,$$

and we define the first summation to be identically zero if n takes the value one.

3.3.iv Equilibrium distribution

If $q=\mu_2-\lambda_2>0$, the dummy variable z_2 ($|z_2|<1$) can always be chosen to satisfy

$$0<\lambda_2 P(z_2)<1 .$$

For such a suitable z_2 we may therefore bound the integral (3.18) by

$$0<J_1(t)<\int_0^t Pe^{(r-q)s}ds = P(r-q)^{-1}(e^{(r-q)t}-1) .$$

Thus

$$0<e^{-rt}J_1(t)<P(r-q)^{-1}(e^{-qt}-e^{-rt}) .$$

Now $r=\mu_1+\nu_1>0$, unless $\mu_1=\nu_1=0$ when the process is trivial. Hence if $q=\mu_2-\lambda_2>0$ we have

$$e^{-rt}J_1(t)\to 0 \quad \text{as} \quad t\to\infty ,$$

and on letting $t\to\infty$ in (3.19) we obtain the equilibrium solution

$$(3.26)\quad G(z_1,z_2;\infty) = \left[\frac{\lambda_2 z_2-\mu_2}{\lambda_2-\mu_2}\right]^{-[\alpha_1\nu_1/(\mu_1+\nu_1)+\alpha_2]/\lambda_2} \times \exp\{[\alpha_1/(\mu_1+\nu_1)](z_1-1)\} .$$

We have previously seen that the p.g.f. $G(z_1,z_2;t)$ given by (3.19) may be partitioned into two separate p.g.f.'s $G_1(z_1;t)$ and $G_2(z_2;t)$. These functions may be thought of as representing two independent processes. However, although $G_1(z_1;t)$ corresponds to the p.g.f. of a simple population process the p.g.f. $G_2(z_2;t)$ does not. In equilibrium this disparity no longer holds. For it follows from (3.26) that $G_1(z_1;\infty)$ corresponds to a simple immigration-death process with parameters α_1 and $\mu_1+\nu_1$, whilst $G_2(z_2;\infty)$ corresponds to a simple immigration-birth-death process with parameters $\alpha_2+\alpha_1\nu_1/(\mu_1+\nu_1)$, λ_2 and μ_2 , respectively.

3.4 Model 3 ($\lambda_2 = \nu_2 = 0$)

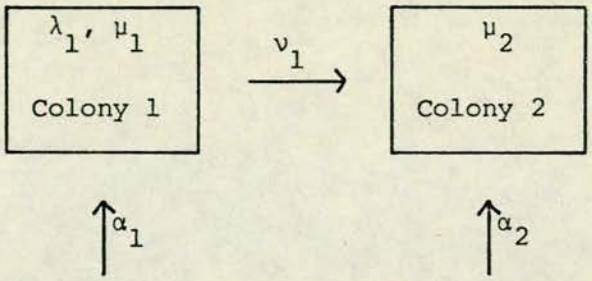
3.4.i Probability generating function

In the previous section we made the assumption that migration could only take place from colony 1 to colony 2, and we also placed $\lambda_1=0$. We shall now retain the assumption of one-way migration, but this time we place $\lambda_2=0$ instead of $\lambda_1=0$. Thus colony 1 develops as a simple immigration-birth-death process with death rate $(\mu_1+\nu_1)$, and colony 2 develops as an immigration-death process having a time-dependent immigration structure. This model is illustrated in figure 3.4, and we shall see that it is considerably more difficult to analyse than the previous two processes.

Under the above restrictions the first two of the auxiliary equations (2.3) become

Figure 3.4

Two-colony process with $\lambda_2=\nu_2=0$



(4.1) $$dz_1/dt = -(\lambda_1 z_1-\mu_1)(z_1-1) - \nu_1(z_2-z_1)$$

(4.2) $$dz_2/dt = -\mu_2(1-z_2) \quad .$$

Equation (4.2) integrates directly to give

(4.3) $$(1-z_2)e^{-\mu_2 t} = A \quad ,$$

where A is a constant of integration, and substitution of (z_2-1) from (4.3) into (4.1) yields the Riccati differential equation

(4.4) $$dz_1/dt = -(\lambda_1 z_1-\mu_1)(z_1-1) - \nu_1(1-z_1) + \nu_1 A e^{\mu_2 t} \quad .$$

Write

$$y_i = 1-z_i \quad (i=1,2) \quad \text{and} \quad \xi_1 = \lambda_1-\mu_1-\nu_1 \quad .$$

Equation (4.4) then becomes

(4.5) $$dy_1/dt = \lambda_1 y_1^2 - \xi_1 y_1 - \nu_1 A e^{\mu_2 t} \quad ,$$

and a further substitution

(4.6) $$x = y_1 - (\xi_1/2\lambda_1)$$

yields

$$dx/dt = \lambda_1 x^2 - [\,(\xi_1^2/4\lambda_1) + \nu_1 A e^{\mu_2 t}\,] \quad . \tag{4.7}$$

Now this equation may be written in the form

$$\frac{dx}{dt} = \frac{\phi'(t)}{h(t)}\, x^2 - \frac{h'(t)}{\phi(t)} \quad , \tag{4.8}$$

say, where the functions $\phi(t)$ and $h(t)$ satisfy the relations

$$\phi'(t)/h(t) = \lambda_1 \tag{4.9}$$

and

$$h'(t)/\phi(t) = (\xi_1^2/4\lambda_1) + \nu_1 A e^{\mu_2 t} \quad . \tag{4.10}$$

Here a dash (') denotes differentiation with respect to t. Once (4.9) and (4.10) are solved for $\phi(t)$ and $h(t)$, equation (4.8) yields the solution

$$x(t) = \frac{-h(t)}{\phi(t)} + \frac{1}{\phi^2(t)} \left[D - \int^t \frac{\phi'(s)ds}{\phi^2(s)h(s)} \right]^{-1} \tag{4.11}$$

where D is a constant of integration (Kamke (1959), p.298, equation (1.33)). This result is easily verified by differentiating both sides of (4.11) with respect to t, and showing that the right hand side agrees with (4.7).

Eliminating $h(t)$ from equations (4.9) and (4.10) we have

$$\phi''(t) = \phi(t)[\,(\xi_1^2/4) + \nu_1 \lambda_1 A e^{\mu_2 t}\,] \quad ,$$

and, with the change of variable

$$r = (2/\mu_2)(\lambda_1 \nu_1 A)^{\frac{1}{2}} e^{\mu_2 t/2} \quad , \tag{4.12}$$

this may be written as

$$r^2\frac{d^2\phi}{dr^2} + r\frac{d\phi}{dr} - \phi[r^2+(\xi_1/\mu_2)^2] = 0 \quad . \tag{4.13}$$

Expression (4.13) is a standard differential equation having the solution

$$\phi(r) = I_{\xi_1/\mu_2}(r)$$

where $I_{\xi_1/\mu_2}(r)$ is the modified Bessel function of the first kind (Abramowitz and Stegun (1965), result (9.6.1)). Replacing r by (4.12) we get

$$\phi(t) = I_{\xi_1/\mu_2}[(2/\mu_2)(\lambda_1\nu_1 A)^{\frac{1}{2}}e^{\mu_2 t/2}] \quad , \tag{4.14}$$

and the substitution of (4.14) into (4.9) gives

$$h(t) = \lambda_1^{-1}\frac{d}{dt}\{I_{\xi_1/\mu_2}[(2/\mu_2)(\lambda_1\nu_1 A)^{\frac{1}{2}}e^{\mu_2 t/2}]\} \quad . \tag{4.15}$$

The substitution of $\phi(t)$ and $h(t)$ from (4.14) and (4.15) into (4.11) gives $x(t)$, whence use of (4.6), namely

$$y_1(t) = x(t) + (\xi_1/2\lambda_1) \quad ,$$

yields

$$y_1(t) = (\xi_1/2\lambda_1) - \phi'(t)/(\lambda_1\phi(t)) + \phi^{-2}(t)\{D-\lambda_1\int^t\phi^{-2}(s)ds\}^{-1} \tag{4.16}$$

where $\phi(t)$ is given by (4.14).

On substituting for $z_1=1-y_1$ and z_2 from (4.16) and (4.3) respectively into the third of the auxiliary equations (2.3), namely

$$\frac{d}{dt}(\log G) = \alpha_1(z_1-1) + \alpha_2(z_2-1) \quad ,$$

and integrating, we have

$$G(z_1,z_2;t) = \text{const.}[\phi(t)]^{\alpha_1/\lambda_1} \exp\{-(\alpha_2 A/\mu_2)e^{\mu_2 t} - (\alpha_1\xi_1/2\lambda_1)t$$

(4.17)

$$- \alpha_1 \int^t \phi^{-2}(s)\{D - \lambda_1 \int^s \phi^{-2}(\tau)d\tau\}^{-1} ds\}.$$

Thus the general solution for $G(z_1,z_2;t)$ may be written in the form

(4.18) $$G(z_1,z_2;t) = f(A,D)[\phi(t)]^{\alpha_1/\lambda_1} \exp\{-(\alpha_2 A/\mu_2)e^{\mu_2 t} - (\alpha_1\xi_1/2\lambda_1)t$$

$$- \alpha_1 \int^t \phi^{-2}(s)\{D - \lambda_1 \int^s \phi^{-2}(\tau)d\tau\}^{-1} ds\}$$

where f is an arbitrary function.

For algebraic simplicity let us assume that both colonies are initially empty. Then at time $t=0$

$$G(z_1,z_2;0) = 1$$

and so (4.18) becomes

$$1 = f(A,D)[\phi(0)]^{\alpha_1/\lambda_1} \exp\{-(\alpha_2 A/\mu_2) - \alpha_1 \int^0 \phi^{-2}(s)\{D-\lambda_1 \int^s \phi^{-2}(\tau)d\tau\}^{-1}ds\}$$

whence

$$G(z_1,z_2;t) = [\phi(t)/\phi(0)]^{\alpha_1/\lambda_1} \exp\{-(\alpha_2 A/\mu_2)(e^{\mu_2 t}-1) - (\alpha_1\xi_1/2\lambda_1)t$$

(4.19)

$$- \alpha_1 \int_0^t \phi^{-2}(s)\{D - \lambda_1 \int^s \phi^{-2}(\tau)d\tau\}^{-1}ds\} .$$

We now have to replace the constants A and D in expression (4.19) by the appropriate functions of z_1, z_2 and t. From (4.3) we have for A,

(4.3) $$A = (1-z_2)e^{-\mu_2 t} ,$$

and so the function

$$\phi(s) = I_{\xi_1/\mu_2}[(2/\mu_2)(\lambda_1\nu_1 A)^{\frac{1}{2}}e^{\mu_2 s/2}] ,$$

which is given by (4.14), is replaced by

$$\psi(s,t;z_2) = I_{\xi_1/\mu_2}\left[(2/\mu_2)(\lambda_1\nu_1(1-z_2))^{\frac{1}{2}}e^{\mu_2(s-t)/2}\right] \tag{4.20}$$

and $\phi(t)$ by

$$\psi(t,t;z_2) = I_{\xi_1/\mu_2}\left[(2/\mu_2)(\lambda_1\nu_1(1-z_2))^{\frac{1}{2}}\right] \equiv \ell(z_2) \quad \text{(say)}. \tag{4.21}$$

If we denote

$$q = (2/\mu_2)(\lambda_1\nu_1(1-z_2))^{\frac{1}{2}}$$

and

$$q(s,t) = qe^{\mu_2(s-t)/2}$$

$d\phi(s)/ds$ becomes

$$\frac{d\psi(s,t;z_2)}{ds} = \tfrac{1}{2}\mu_2 q(s,t)\,\frac{d}{dq(s,t)}\{I_{\xi_1/\mu_2}[q(s,t)]\}\ .$$

In particular, $d\phi(s)/ds$ at $s=t$ becomes

$$\left.\frac{d\psi(s,t;z_2)}{ds}\right|_{s=t} = \tfrac{1}{2}\mu_2 q\,\frac{d}{dq}\{I_{\xi_1/\mu_2}(q)\} \equiv n(z_2) \quad \text{(say)}. \tag{4.22}$$

It now follows from (4.16) and (4.20)-(4.22) that

$$D = \lambda_1\int^t \psi^{-2}(\tau,t;z_2)d\tau + H(z_1,z_2) \equiv N(z_1,z_2) \quad \text{(say)} \tag{4.23}$$

where

$$H(z_1,z_2) = \ell^{-2}(z_2)[(1-z_1) - (\xi_1/2\lambda_1) + (n(z_2)/\lambda_1\ell(z_2))]^{-1}\ . \tag{4.24}$$

Note that the function $N(z_1,z_2)$ is independent of t. For on putting $s=t-\tau$ we get

$$(4.25)\qquad \int^{t}\psi^{-2}(\tau,t;z_2)d\tau = -\int^{o}I^{-2}_{\xi_1/\mu_2}[qe^{-\mu_2 s/2}]ds \quad ,$$

and this integral is clearly independent of t .

Thus the integral in expression (4.19), namely

$$R(z_1,z_2;t) \equiv \int_o^t \phi^{-2}(s)\{D - \lambda_1\int^{s}\phi^{-2}(\tau)d\tau\}^{-1}ds \quad ,$$

may be written in the form

$$(4.26)\quad R(z_1,z_2;t) = \int_o^t I^{-2}_{\xi_1/\mu_2}[qe^{\mu_2(s-t)/2}]\{H(z_1,z_2) + \lambda_1\int_s^t I^{-2}_{\xi_1/\mu_2}[qe^{\mu_2(\tau-t)/2}]d\tau\}^{-1}ds \quad .$$

Denote the indefinite integral $L(t)$ by

$$(4.27)\qquad L(t) = \int^{t} I^{-2}_{\xi_1/\mu_2}[qe^{-\mu_2 s/2}]ds \quad .$$

Then (4.26) becomes

$$R(z_1,z_2;t) = \int_o^t I^{-2}_{\xi_1/\mu_2}[qe^{\mu_2(s-t)/2}]\{H(z_1,z_2) + \lambda_1[L(t-s)-L(0)]\}^{-1}ds \quad ,$$

and on putting $y=t-s$ we have

$$R(z_1,z_2;t) = \int_o^t \frac{dL(y)}{dy}\{H(z_1,z_2) + \lambda_1[L(y)-L(0)]\}^{-1}dy \quad .$$

This expression integrates to give

$$(4.28)\quad R(z_1,z_2;t) = \lambda_1^{-1}\log\{[H(z_1,z_2) + \lambda_1(L(t)-L(0))]/H(z_1,z_2)\} \quad .$$

Define the definite integral $L_1(t)$ by

$$(4.29)\qquad L_1(t) \equiv L(t) - L(0) \quad .$$

The solution (4.19) may now be written in the form

$$G(z_1,z_2;t) = \left[\frac{\psi(t,t;z_2)}{\psi(0,t;z_2)} \cdot \frac{H(z_1,z_2)}{H(z_1,z_2)+\lambda_1 L_1(t)}\right]^{\alpha_1/\lambda_1}$$

(4.30)

$$\times \exp\{-(\alpha_2/\mu_2)(1-z_2)(1-e^{-\mu_2 t}) - (\alpha_1\xi_1/2\lambda_1)t\} \quad .$$

Expression (4.30) was verified by back-substitution into equation (2.2.2), and we shall later show that $G(1,1;t) \equiv 1$ for all $t \geq 0$ as required.

Placing $t=0$ in (4.30) we obtain $G(z_1,z_2;0) = 1$, so ensuring that the initial conditions are satisfied, whilst placing $\alpha_1=0$ we get the standard p.g.f. for the simple immigration-death process (Cox and Miller (1965), p.168). In this latter case, colony 1 remains permanently empty. Puri (1968) obtains an expression equivalent to that of (4.30) for the case in which immigration is excluded $(\alpha_1=\alpha_2=0)$. He assumes the initial condition $X_1(0)=1$, $X_2(0)=0$, whilst we take $X_1(0) = X_2(0) = 0$, and so his result cannot be derived from (4.30).

3.4.ii Equilibrium distribution

The rate of departure of an individual from colony 1, either through death or migration to colony 2, is equal to $\mu_1 + \nu_1$. Thus if $\lambda_1 < \mu_1 + \nu_1$, i.e. $\xi_1 < 0$, we may expect a limiting equilibrium population size distribution to develop as $t \to \infty$. Now both $\psi(t,t;z_2)$ and $H(z_1,z_2)$ in expression (4.30) are independent of t, and so let us first evaluate

$$Y \equiv \lim_{t\to\infty}\{\psi(0,t;z_2)\,[H(z_1,z_2) + \lambda_1 L_1(t)]\}^{\alpha_1/\lambda_1} e^{(\alpha_1\xi_1/2\lambda_1)t}$$

which, on using (4.20) and (4.29), we may show becomes

$$(4.31)\quad Y^{\lambda_1/\alpha_1} = \lim_{t\to\infty} e^{\xi_1 t/2}\{I_{\xi_1/\mu_2}(qe^{-\mu_2 t/2})\}[H(z_1,z_2) + \lambda_1\int_o^t I_{\xi_1/\mu_2}(qe^{-\mu_2 s/2})ds] .$$

In order to evaluate this limit we use the following asymptotic result. When x is small and p is fixed

$$(4.32)\qquad I_p(x) = \{(\tfrac{1}{2}x)^p/\Gamma(p+1)\}[1+O(x^2)] .$$

(Abramowitz and Stegun (1965), result (9.6.10)). Here $\Gamma(p+1)$ denotes the Gamma function.

Result (4.32) is only valid when $p\neq -1,-2,\ldots,$ and so we have to consider the following two cases.

(i) If $\xi_1/\mu_2 \neq -1,-2,\ldots,$ expression (4.31) gives

$$Y^{\lambda_1/\alpha_1} = \lim_{t\to\infty} e^{\xi_1 t/2}(\tfrac{1}{2}qe^{-\mu_2 t/2})^{\xi_1/\mu_2}\Gamma^{-1}(1+\xi_1/\mu_2)[H(z_1,z_2) + \lambda_1\int_o^t I_{\xi_1/\mu_2}(qe^{-\mu_2 s/2})ds]$$

which, on using (4.29), we may show to be

$$(4.33)\qquad = (\tfrac{1}{2}q)^{\xi_1/\mu_2}\Gamma^{-1}(1+\xi_1/\mu_2)[H(z_1,z_2) + \lambda_1 L_1(\infty)] .$$

Combining (4.21), (4.30) and (4.33) we obtain the solution

$$(4.34)\qquad G(z_1,z_2;\infty) = \exp\{(\alpha_2/\mu_2)(z_2-1)\}\left[\frac{I_{\xi_1/\mu_2}(q)H(z_1,z_2)\Gamma(1+\xi_1/\mu_2)}{(\tfrac{1}{2}q)^{\xi_1/\mu_2}[H(z_1,z_2)+\lambda_1 L_1(\infty)]}\right]^{\alpha_1/\lambda_1} .$$

(ii) Suppose $\xi_1/\mu_2 = -p$ for $p=1,2,\ldots,$ and note that for integer p

$$I_{-p}(x) \equiv I_p(x) \quad .$$

If we use result (4.32) expression (4.31) may be written as

$$Y^{\lambda_1/\alpha_1} = \lim_{t\to\infty} e^{\xi_1 t/2}(\tfrac{1}{2}qe^{-\mu_2 t/2})^p(p!)^{-1}[H(z_1,z_2)+\lambda_1\int_o^t I_p^{-2}(qe^{-\mu_2 s/2})ds]$$

which, on putting

$$x(s) = qe^{-\mu_2 s/2} \quad ,$$

we may show to be

$$(4.35) = \lim_{t\to\infty} (\tfrac{1}{2}q)^p(p!)^{-1}e^{\xi_1 t}[H(z_1,z_2)-2(\lambda_1/\mu_2)\int_{x(0)}^{x(t)} x^{-1}I_p^{-2}(x)dx] \quad .$$

But from (4.32)

$$\int_{x(0)}^{x(t)} x^{-1}I_p^{-2}(x)dx = \int_{x(0)}^{x(t)} x^{-1}(\tfrac{1}{2}x)^{-2p}[(p!)^{-1}+O(x^2)]^{-2}dx$$

$$= \int_{x(0)}^{x(t)} 2^{2p}(p!)^2 x^{-(2p+1)}[1+O(x^2)]dx$$

$$(4.36) \qquad \sim -2^{2p}(p!)^2(1/2p)q^{-2p}e^{-\xi_1 t}$$

where $p=1,2,\ldots,$ and $\xi_1 < 0$. Placing (4.36) in (4.35) we get

$$Y^{\lambda_1/\alpha_1} = (\lambda_1/\mu_2)(\tfrac{1}{2}q)^{-p}(p-1)!$$

whence it follows from (4.21) and (4.30) that

$$(4.37) \quad G(z_1,z_2;\infty) = \exp\{(\alpha_2/\mu_2)(z_2-1)\}$$

$$\times [(\mu_2/\lambda_1)I_p(q)H(z_1,z_2)(\tfrac{1}{2}q)^p/(p-1)!]^{\alpha_1/\lambda_1} \quad .$$

Results (4.34) and (4.37) were shown to satisfy the differential equation (2.2.2) by back-substitution.

3.4.iii Marginal distributions

As migrants may travel only from colony 1 to colony 2, we expect that the marginal distribution of colony 1 is that of a simple immigration-birth-death process with parameters α_1, λ_1 and $\mu_1+\nu_1$ respectively. However, immigration to colony 2 is composed of a simple immigration process with rate α_2 and a more complicated component due to the migration of individuals to colony 2 from colony 1. Thus we may reasonably expect the marginal distribution of colony 2 to be relatively complex compared with that of colony 1. This is precisely what happens. For we see that if we place $z_1=1$ in (4.30) we obtain virtually no simplification. In fact the only alteration to $G(z_1,z_2;t)$ is the replacement of $H(z_1,z_2)$ by $H(1,z_2)$, where from (4.24)

$$H(1,z_2) = [n(z_2) - \tfrac{1}{2}\xi_1 \ell(z_2)]/\lambda_1 \ell^3(z_2) \quad .$$

The marginal equilibrium p.g.f. $G(1,z_2;\infty)$, obtained by placing $z_1=1$ in (4.34) and (4.37), is similarly affected. On the other hand, if we put $z_2=1$ in expression (4.30) we obtain

$$(4.38)\quad G(z_1,1;t) = \xi_1^{\alpha_1/\lambda_1}[\lambda_1 e^{\xi_1 t} - (\mu_1+\nu_1) - \lambda_1 z_1(e^{\xi_1 t}-1)]^{-\alpha_1/\lambda_1} \quad ,$$

which is the p.g.f. of a simple immigration-birth-death process with parameters α_1,λ_1 and $\mu_1+\nu_1$ respectively (see Bartlett (1966), p.81). The proof of result (4.38) is omitted from this thesis as it involves a lengthy piece of algebra.

We have previously shown that expression (4.30) for the p.g.f. $G(z_1,z_2;t)$ satisfies the differential equation (2.2.2) and the initial condition

$$G(z_1,z_2;0) \equiv 1 \quad .$$

If we place $z_1=1$ in (4.38) we see that $G(z_1,z_2;t)$ satisfies the further requirement that $G(1,1;t) \equiv 1$ for all $t \geq 0$.

3.5 Recursive Solutions to the Kolmogorov Forward Equation

3.5.i Recursive solutions

In the literature, probabilities which arise in stochastic models are nearly always determined by the solution of a differential equation for a relevant generating function. Although the derivation of a solution for such an equation can be mathematically satisfying, the solution obtained will often be too complex to allow the extraction of the probabilities themselves. Expression (4.30) of the previous section provides an example.

Severo (1967) partly overcame this difficulty by developing a method for dealing with the original system of ordinary differential-difference equations for the probabilities. He presents an iterative solution to the vector-matrix differential equation

(5.1) $$d\underline{x}(t)/dt = B\underline{x}(t)$$

where B is a special type of triangular matrix. Thus we have the solution to any model whose original set of equations, which describe the probabilities, can be written in the form (5.1). Note that this method is iterative in the sense that at each stage of the procedure one uses some of the results of the previous steps in order to obtain entries in the solution matrix.

Severo gives as examples the solutions to the simple and the

general stochastic epidemic. Let $p_{rs}(t)$ denote the probability that at time $t \geq 0$ there are r susceptibles and s infectives given that $r+s=N$ at time $t=0$. Then on making the transformation

$$p_{rs}(t) = x_k(t)$$

where

$$\text{(5.2)} \quad \begin{array}{lll} \text{(i)} & r+s=N \quad \text{and} & k = N-r+1 \qquad \text{(simple epidemic)} \\ \text{(ii)} & 0 \leq r+s \leq N \quad \text{and} & k = \tfrac{1}{2}(N+1)(N+2) - (N+1)r - s + \tfrac{1}{2}(r-1)r \\ & & \qquad \text{(general epidemic)} \end{array}$$

we find that the resulting system of ordinary differential equations for $\{x_k(t)\}$ is of the required form (5.1).

Severo (1969a) generalizes this approach with a recursion theorem that leads to the solution for any finite system of differential-difference equations with constant triangular coefficient matrix. He illustrates this method of solution by discussing its applications to families of bounded multi-dimensional pure birth and pure death processes, stochastic cross-infection among several otherwise isolated groups, and to generalizations of the general stochastic epidemic. This latter problem was later treated more extensively by Severo (1969b) in an analysis of Downton's (1968) carrier-borne model. This last model may be interpreted as a generalization of the general stochastic epidemic allowing for the removal of both susceptibles and infectives.

Billard (1973) considers the general stochastic epidemic and notes that in none of the previous methods, including Severo's, is the actual nature of the solution apparent. She uses an

appropriate partition of the coefficient matrix to develop solutions which indicate their form very clearly. Here the total population size N is bounded, but she remarks that because the special nature of the solution is now revealed, asymptotic solutions for large values of N could be found.

In each of the three cases so far examined in this chapter, births are permitted to occur in at most one of the colonies. When both $\lambda_1 > 0$ and $\lambda_2 > 0$ the auxiliary equations (2.3) yield two simultaneous Riccati differential equations in terms of z_1, z_2 and t, and the probability generating function approach clearly becomes far too complex to be worth pursuing. It seems natural to ask whether the Severo approach may be used instead.

One difficulty is that Severo and Billard work with finite systems of differential equations whilst the system (2.2.1) is infinite. In principle this difficulty may be overcome by constraining the total population size to be at most N_0 (say), where $0 < N_0 < \infty$. To achieve this we put

$$\Pr[\text{a birth or immigration occurs in } (t,t+\delta t) \mid X_1(t)+X_2(t)=N_0] = 0 .$$

The problem then reduces to a search for a suitable transformation from $p_{ij}(t)$ to $x_k(t)$ similar to (5.2). However, as $0 \leq X_1(t)+X_2(t) \leq N_0$ for all $t \geq 0$ this approach would yield an approximate and not an exact solution. We shall therefore delay an examination of this approach until the next chapter.

Returning to the unrestricted situation let us parallel Severo's technique by writing the differential equations (2.2.1) in triangular form. This may be achieved by disregarding death and

imposing one-way migration. For if we denote the vector

$$\underline{p}(t) = (p_{10}(t), p_{01}(t), \ldots, p_{n0}(t), p_{n-1,1}(t), \ldots, p_{0n}(t), p_{n+1,0}(t), \ldots)' \tag{5.3}$$

and put

$$\mu_1 = \mu_2 = \nu_2 = 0 \quad ,$$

we may write the differential-difference equations (2.2.1) in the vector-matrix form

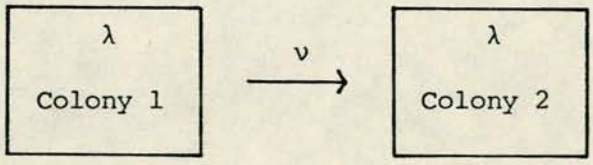
$$d\underline{p}(t)/dt = B\underline{p}(t) \tag{5.4}$$

where B is a lower-triangular matrix independent of t . The probabilities $p_{ij}(t)$ may now be determined in the order written in the vector (5.3). Note that as $\nu_2=0$, colony 1 develops as a simple immigration-birth-death process with parameters α_1, λ_1 and ν_1 respectively.

For algebraic convenience let us place $\lambda_1=\lambda_2=\lambda$; $\alpha_1=\alpha_2=0$ and $\nu_1=\nu$ (this particular model is illustrated in figure 3.5). Equations (2.2.1) then become

Figure 3.5

Two-colony process with $\lambda_1=\lambda_2=\lambda$; $\mu_1=\mu_2=\nu_2=\alpha_1=\alpha_2=0$; $\nu_1=\nu$



$$dp_{n0}(t)/dt = (n-1)\lambda p_{n-1,0}(t) - n(\lambda+\nu)p_{n0}(t) \tag{5.5}$$

$$dp_{nj}(t)/dt = (n-1)\lambda p_{n-1,j}(t) + (j-1)\lambda p_{n,j-1}(t) + (n+1)\nu p_{n+1,j-1}(t) - [n(\lambda+\nu)+j\lambda]p_{nj}(t) \qquad (j>0) \; . \tag{5.6}$$

Suppose that we have the initial conditions

$$p_{ij}(0) = \begin{cases} 1 : i=1, j=0 \\ 0 : \text{otherwise} \end{cases} . \tag{5.7}$$

Then by using either a single induction argument or a generating function approach, we may integrate equation (5.5) and so derive the probabilities

$$p_{n0}(t) = [\lambda/(\lambda+\nu)]^{n-1} e^{-(\lambda+\nu)t}[1-e^{-(\lambda+\nu)t}]^{n-1} \qquad (n \geq 1) . \tag{5.8}$$

Successive use of equation (5.6) with $j=1,2,\ldots,$ will now generate the sequences $\{p_{n1}(t)\}, \{p_{n2}(t)\},\ldots$.

The algebraic complexity of this section will probably convince the reader that this approach, though yielding the correct solution, does not really provide an acceptable way of generating the probabilities $p_{ij}(t)$. Nevertheless, the technique itself is of interest and it is primarily for this reason that we include the following analysis in the thesis. Moreover, the situation in which $\lambda_1=\lambda_2=\lambda$, $\nu_1=\nu$ and all other transition rates are equal to zero is the simplest spatial model possible for which both birth rates are non-zero. The computational difficulty we shall experience for this particular model will illustrate the immense severity of the general problem. Note that the auxiliary equations (2.3) yield

$$G = A$$

$$z_2 = (1-Be^{-\lambda t})^{-1} \tag{5.9}$$

$$dz_1/dt = -\lambda z_1^2 + (\lambda+\nu)z_1 - \nu z_2 \tag{5.10}$$

where A and B are constants of integration. Substituting for

z_2 from (5.9) into (5.10) we get an equation of Riccati form similar to (4.4), and the solution of (5.10) is even more complex than that of (4.4).

Assume that the probabilities $p_{nj}(t)$ are known for all $n=0,1,2,\ldots,$ and $j=0,1,\ldots,m-1$. Then equation (5.6) gives a set of first-order linear differential equations for the unknown $p_{nm}(t)$ $(n \geq 0)$ in terms of known $p_{nj}(t)$ $(0 \leq j < m,\ n \geq 0)$. Denoting the vector

$$\underline{p}_m(t) = (\ldots\ldots, p_{n0}(t), p_{n-1,1}(t), \ldots, p_{n-m,m}(t), p_{n+1,0}(t), \ldots\ldots)' ,$$

we see that equations (5.6) may be expressed in the vector-matrix form

$$d\underline{p}_m(t)/dt = B_m \underline{p}_m(t) \tag{5.11}$$

where B_m is a lower-triangular matrix with diagonal elements in the following order.

$$\begin{array}{l} -(\lambda+\nu)\ ,\ -\lambda \\ -2(\lambda+\nu)\ ,\ -(\lambda+\nu)-\lambda\ ,\ -2\lambda \\ \qquad\qquad \ldots\ldots\ldots \\ -r(\lambda+\nu)\ ,\ -(r-1)(\lambda+\nu)-\lambda\ ,\ -(r-2)(\lambda+\nu)-2\lambda\ ,\ \ldots\ ,\ -d_r-[\min(r,m)]\lambda \\ \qquad\qquad \ldots\ldots\ldots \end{array} \tag{5.12}$$

where $d_r=0$ if $r \leq m$ and $d_r=(r-m)(\lambda+\nu)$ if $r > m$. Here r takes the values $1,2,3,\ldots,$ and the elements of (5.12) are read from left to right row by row. The representation (5.11), (5.12) is obtained if we ignore all equations (5.6) which contain $dp_{nj}(t)/dt$ with $j > m$ on their left hand side.

The general solution to equations of the lower-triangular form (5.11) (Severo (1969a), theorem 1) may not be used here as B_m is

an infinite matrix. However, as B_m is of lower-triangular form, with diagonal elements given by (5.12), it follows that for some constants $c_{nm}(r,i)$ we may express the solution of (5.11) in the form

$$p_{nm}(t) = \sum_{r=0}^{m} e^{-r\lambda t} \sum_{i=0}^{n+m-r} c_{nm}(r,i) e^{-i(\lambda+\nu)t} \tag{5.13}$$

where

$$c_{nm}(0,0) = 0 \qquad (n,m=0,1,2,\ldots) .$$

Substituting (5.13) into equations (5.6) with $j=m$, comparing coefficients of $\exp\{-t[r\lambda+i(\lambda+\nu)]\}$ on both sides of the resulting expressions, and denoting

$$c_{hk}(r,i) \equiv 0 \qquad (h<0 \text{ or } k<0) ,$$

we get the following two sets of equations:

(i) for $r=0,1,\ldots,m-1$

$$c_{nm}(r,i)[(m-r)\lambda+(n-i)(\lambda+\nu)] = (n-1)\lambda c_{n-1,m}(r,i) \tag{5.14}$$

$$+ (m-1)\lambda c_{n,m-1}(r,i) + (n+1)\nu c_{n+1,m-1}(r,i) \qquad (0 \le i < n+m-r) ,$$

$$-(m-r)c_{nm}(r,n+m-r) = (n+1)c_{n+1,m-1}(r,n+m-r) \qquad (i=n+m-r) \tag{5.15}$$

whilst

(ii) for $r=m$

$$(n-i)(\lambda+\nu)c_{nm}(m,i) = (n-1)\lambda c_{n-1,m}(m,i) \qquad (0 \le i < n) ; \tag{5.16}$$

the equation for $r=m$, $i=n$ is automatically satisfied. Thus our problem reduces to a search for coefficients $c_{nm}(r,i)$ which satisfy equations (5.14)-(5.16).

Solving equations (5.16) recursively in n, and denoting

$a=\lambda/(\lambda+\nu)$, we obtain

$$c_{nm}(m,i) = \binom{n-1}{i-1} a^{n-i} c_{im}(m,i) \qquad (0 < i < n) , \tag{5.17}$$

whilst if $i=0$

$$c_{nm}(m,0) = 0 \qquad (n > 0) \tag{5.18}$$

with $c_{0m}(m,0)$ as yet undefined.

The coefficients $c_{nm}(r,i)$ in the representation (5.13) are defined only for $i=0,1,\ldots,n+m-r$. Let us denote

$$c_{nj}(r,i) = 0 \qquad (i > n+j-r) .$$

Then (5.15) is equivalent to (5.14) when $i=n+m-r$. Writing equations (5.14) in terms of the generating function

$$H_{ri}^{(m)}(z) = \sum_{n=0}^{\infty} c_{nm}(r,i) z^n \tag{5.19}$$

we get

$$(\lambda+\nu) z \frac{dH}{dz}^{(m)} + [\,(m-r)\lambda - i(\lambda+\nu)\,] H^{(m)} = \lambda z^2 \frac{dH}{dz}^{(m)} + (m-1)\lambda H^{(m-1)} + \nu \frac{dH}{dz}^{(m-1)} ,$$

and on putting $q=(m-r)a-i$ we may integrate this differential equation to obtain

$$H_{ri}^{(m)}(z) = \left[\frac{z}{1-az}\right]^{-q} \int^{z} y^{q-1} (1-ay)^{-(q+1)} [\,(m-1)aH_{ri}^{(m-1)}(y) + (1-a)\frac{dH_{ri}^{(m-1)}(y)}{dy}\,] dy . \tag{5.20}$$

As $a=\lambda/(\lambda+\nu)$, q will in general be non-integer. In this general case the constant of integration in (5.20) must be zero otherwise (5.20) would contain the additional term

$$\text{const.} z^{-q} (1-az)^{q} ,$$

and this expands to give non-integer powers of z unless

$$\text{const.} = 0 \quad .$$

Remembering that the $c_{nj}(r,i)$ are assumed known for $j < m$, we see that expressions (5.19) and (5.20) determine

$$c_{nm}(r,i) \qquad (r=0,\ldots,m-1) \quad ,$$

whilst (5.18) shows that

$$c_{nm}(m,0) = 0 \qquad (n > 0)$$

and (5.17) expresses

$$c_{nm}(m,i) \qquad (0 < i < n)$$

in terms of the $c_{im}(m,i)$. It therefore remains to evaluate $c_{im}(m,i)$ for $i \geq 0$.

Placing $t=0$ in the representation (5.13), substituting for $c_{nm}(m,i)$ from (5.17) and (5.18), and assuming the initial conditions

$$p_{ij}(0) = \begin{cases} 1 : i=1,\ j=0 \\ 0 : \text{otherwise} \end{cases} \quad , \tag{5.21}$$

we obtain the following set of equations for $m > 0$:

$$0 = \sum_{r=0}^{m-1} \sum_{i=0}^{m-r} c_{0m}(r,i) + c_{0m}(m,0) \qquad (n=0) \quad , \tag{5.22}$$

$$0 = \sum_{r=0}^{m-1} \sum_{i=0}^{n+m-r} c_{nm}(r,i) + \sum_{i=1}^{n} \binom{n-1}{i-1} a^{n-i} c_{im}(m,i) \qquad (n > 0) \quad . \tag{5.23}$$

Expressions (5.20) and (5.22) determine $c_{0m}(m,0)$, whilst to evaluate the $c_{im}(m,i)$ $(i > 0)$ we multiply both sides of (5.23) by z^n and sum over $n=1,2,\ldots,$ to obtain

$$0 = \sum_{n=1}^{\infty} z^n \sum_{r=0}^{m-1} \sum_{i=0}^{n+m-r} c_{nm}(r,i) + \sum_{n=1}^{\infty} z^n \sum_{i=1}^{n} \binom{n-1}{i-1} a^{n-i} c_{im}(m,i)$$

(5.24)

$$\equiv S_1(z) + S_2(z) \qquad \text{(say)} \quad .$$

Thus

$$-S_1(z) = \sum_{i=1}^{\infty} c_{im}(m,i) \left[\frac{z}{1-az}\right]^i \quad .$$

Writing

$$y = z/(1-az)$$

we get

(5.25)
$$\sum_{i=1}^{\infty} c_{im}(m,i) y^i = -S_1\left[\frac{y}{1+ay}\right] \quad .$$

But we have already determined $S_1(z)$ as an infinite polynomial in z . Thus the $c_{im}(m,i)$ $(i > 0)$ may be obtained by expanding $S_1[y/(1+ay)]$ as an infinite polynomial in y and comparing coefficients of y^i on both sides of (5.25). The $c_{nm}(m,i)$ $(0 < i < n)$ may now be obtained from (5.17).

Thus given the values of the $c_{nj}(r,i)$ for $j < m$, the above technique yields the $c_{nm}(r,i)$ $(j=m)$. Hence the solution to equations (5.5) and (5.6) may be obtained recursively as m takes the successive values $m=1,2,3,\ldots$.

3.5.ii <u>An application with $m=1$</u>

As we have already shown that

(5.8) $$p_{n0}(t) = [\lambda/(\lambda+\nu)]^{n-1} e^{-(\lambda+\nu)t} [1-e^{-(\lambda+\nu)t}]^{n-1} \qquad (n \geq 1) \ ,$$

let us illustrate the above technique by finding the probabilities $\{p_{n1}(t)\}$ $(n \geq 0)$. To simplify the notation denote

$$a_{ni} = c_{n1}(0,i) \quad \text{and} \quad b_{ni} = c_{n1}(1,i) \ ,$$

so that (5.13) becomes

$$p_{n1}(t) = \sum_{i=0}^{n+1} a_{ni} e^{-i(\lambda+\nu)t} + e^{-\lambda t} \sum_{i=0}^{n} b_{ni} e^{-i(\lambda+\nu)t} \ . \tag{5.26}$$

Expressions (5.17) and (5.18) give

$$b_{ni} = \binom{n-1}{i-1} a^{n-i} b_{ii} \qquad (0 < i < n) \tag{5.27}$$

with

$$b_{n0} = 0 \qquad (n > 0) \ . \tag{5.28}$$

To evaluate the a_{ni} we first write (5.8) in the form

$$p_{n0}(t) = \sum_{i=0}^{n-1} a^{n-1} \binom{n-1}{i} (-1)^i e^{-(i+1)(\lambda+\nu)t}$$

so that

$$c_{n0}(0,i) = \binom{n-1}{i-1} a^{n-1} (-1)^{i-1} \qquad (1 \leq i \leq n)$$

$$c_{00}(0,i) = 0 \ .$$

Thus from (5.19)

$$H_{0i}^{(0)}(z) = \sum_{n=0}^{\infty} c_{n0}(0,i) z^n$$

$$= z(-az)^{i-1}(1-az)^{-i}$$

which, on using (5.20), yields

$$H_{0i}^{(1)} \equiv \sum_{n=0}^{\infty} a_{ni} z^n = i(1-a)(-a)^{i-1}\left[\left[\frac{1}{a-1}\right]\left[\frac{z}{1-az}\right]^{i-1} + 2\left[\frac{z}{1-az}\right]^{i} + \left[\frac{a^2}{a+1}\right]\left[\frac{z}{1-az}\right]^{i+1}\right] \ . \tag{5.29}$$

Note that from (5.26) $a_{ni}=0$ for $n < i-1$, so a comparison of coefficients of z^n on both sides of (5.29) gives

$$a_{i-1,i} = ia^{i-1}(-1)^i$$

$$a_{ii} = -i(-a)^{i-1}[a(i+1)-2]$$

$$a_{ni} = i(1-a)(-a)^{i-1}[(a-1)^{-1}\binom{n-1}{i-2}a^{n-i+1} + 2\binom{n-1}{i-1}a^{n-i}$$

(5.30)
$$+ (a+1)^{-1}\binom{n-1}{i}a^{n-i+1}] \qquad (n > i) .$$

It now remains to evaluate the b_{ii} $(i \geq 0)$. From (5.22)

$$\sum_{i=0}^{1} a_{0i} + b_{00} = 0 \qquad (n=0) .$$

But from (5.30)

(5.31)
$$a_{01} = -1 ,$$

and so as $a_{00}=0$ by definition, we have

(5.32)
$$b_{00}=1 .$$

Expression (5.24) yields

$$S_1(z) = \sum_{n=1}^{\infty} z^n \sum_{i=0}^{n+1} a_{ni}$$

$$= \sum_{i=0}^{\infty} \sum_{n=i-1}^{\infty} a_{ni} z^n - a_{01} .$$

Using (5.29) and (5.31) we get

$$S_1(z) = \sum_{i=0}^{\infty} H_{0i}^{(1)}(z) + 1 .$$

Using (5.29) with $y=z/(1-az)$ we get

$$S_1[y/(1+ay)] = 1 + (1-a)[(a-1)^{-1}+2y+a^2(a+1)^{-1}y^2] \sum_{i=0}^{\infty} i(-ay)^{i-1} ,$$

whence on applying (5.25) we have

(5.33)
$$\sum_{i=1}^{\infty} b_{ii} y^i = -1 + (a-1)[(a-1)^{-1}+2y+a^2(a+1)^{-1}y^2] \sum_{i=0}^{\infty} i(-ay)^{i-1} .$$

Comparing coefficients of y^i on both sides of (5.33) we obtain

$$b_{ii} = 2a(a+1)^{-1}(i+a)(-a)^{i-2} \qquad (i \geq 1) \quad . \tag{5.34}$$

We now substitute for the a_{ni} and b_{ni} from (5.27), (5.28), (5.30)-(5.32) and (5.34) into (5.26) and write

$$s = e^{-(\lambda+\nu)t} \quad ,$$

whence

$$p_{01}(t) = e^{-\lambda t}(1-e^{-\nu t}) \qquad (n=0) \tag{5.35}$$

and

$$\begin{aligned} p_{n1}(t) = a^{n-1}s(1-s)^{n-2}\{&-as[s(n+1)-2] + 2(a-1)(ns-1) \\ &- a(a-1)(a+1)^{-1}(n-1) \qquad (n \geq 1) \\ &+ 2e^{-\lambda t}(a+1)^{-1}[(ns-1)-a(1-s)]\} \quad . \end{aligned} \tag{5.36}$$

Expressions (5.35) and (5.36) satisfy the initial conditions $p_{n1}(0)=0$ $(n=0,1,2,\ldots)$, and they were verified by back-substitution into the differential equations (5.5) and (5.6).

CHAPTER 4

APPROXIMATE SOLUTIONS TO THE KOLMOGOROV FORWARD EQUATIONS: TWO-COLONY MODELS

4.1 Introduction

We have so far derived the first- and second-order moments of the process and have obtained exact solutions for the p.g.f. $G(z_1,z_2;t)$ in several special cases. However, whilst the solutions for model 1 $(\lambda_1 = \lambda_2 = 0)$ and model 2 $(\lambda_1 = \nu_2 = 0)$ are manageable, the solution for model 3 $(\lambda_2 = \nu_2 = 0)$ is clearly not; expression (3.4.30) is far too complicated to be of real use. Thus, even if the general solution were available in closed form (which it is not), it would certainly be of an extremely intricate nature. Further attempts to obtain exact expressions for the p.g.f. would seem to be of little avail.

In this Chapter we shall turn our attention to the derivation of approximate solutions to the Kolmogorov forward equations. First I approximate the process itself by modifying the birth mechanism, thereby developing an exact solution for this new process which closely relates to the Poisson distribution. Whilst the first-order moments of the original and modified processes coincide, the second-order moments do not. I shall partly remedy this defect by deriving a bivariate negative-binomial approximation. Lastly I discuss approximations derived by placing an upper bound on the total population size.

4.2 Approximate Solutions based on a Modified Birth Mechanism

4.2.i Probability generating function

Although equation (2.2.2) has so far not yielded an exact solution in closed form, we may obtain an approximate solution by slightly modifying the birth mechanism. Let us take the probability of a birth in colony i in the small time interval $(t,t+\delta t)$ to be $\lambda_i m_i(t)\delta t+o(\delta t)$, where $m_i(t)$ denotes the mean number of individuals in colony i $(i=1,2)$ at time t. Then equations (2.2.1) become modified to

$$\begin{aligned} dp_{ij}(t)/dt &= \lambda_1 m_1(t) p_{i-1,j}(t) + \lambda_2 m_2(t) p_{i,j-1}(t) + \alpha_1 p_{i-1,j}(t) \\ &+ \alpha_2 p_{i,j-1}(t) + \mu_1(i+1) p_{i+1,j}(t) + \mu_2(j+1) p_{i,j+1}(t) \\ &+ \nu_1(i+1) p_{i+1,j-1}(t) + \nu_2(j+1) p_{i-1,j+1}(t) \\ &- p_{ij}(t)[\lambda_1 m_1(t) + \lambda_2 m_2(t) + \alpha_1 + \alpha_2 + i(\mu_1+\nu_1) + j(\mu_2+\nu_2)] \end{aligned} \tag{2.1}$$

where the $p_{ij}(t)$ now represent the probabilities for the approximation. Note that births have effectively been replaced by time-dependent immigration. On multiplying both sides of (2.1) by $z_1^i z_2^j$ and summing over the range $0 < i,j < \infty$, we see that the equation corresponding to (2.2.2) for the p.g.f.

$$G(z_1,z_2;t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij}(t) z_1^i z_2^j$$

is given by

$$\begin{aligned} \frac{\partial G}{\partial t} &= \frac{\partial G}{\partial z_1}[\mu_1(1-z_1)+\nu_1(z_2-z_1)] + \frac{\partial G}{\partial z_2}[\mu_2(1-z_2)+\nu_2(z_1-z_2)] \\ &+ G[(\lambda_1 m_1(t)+\alpha_1)(z_1-1) + (\lambda_2 m_2(t)+\alpha_2)(z_2-1)] \quad . \end{aligned} \tag{2.2}$$

We shall retain the initial condition (2.2.3), namely

$$G(z_1,z_2;0) = z_1^{a_1} z_2^{a_2} \quad . \tag{2.3}$$

Solving (2.2) in the usual manner we first write down the auxiliary equations

$$(2.4)\qquad \frac{dt}{1} = \frac{-dz_1}{\mu_1(1-z_1)+\nu_1(z_2-z_1)} = \frac{-dz_2}{\mu_2(1-z_2)+\nu_2(z_1-z_2)}$$

$$= \frac{dG}{G[(\lambda_1 m_1(t)+\alpha_1)(z_1-1) + (\lambda_2 m_2(t)+\alpha_2)(z_2-1)]} \quad ,$$

and then note that the first three expressions in (2.4) are identical with those of the auxiliary equations (3.2.3) with $\lambda_1 = \lambda_2 = 0$. Thus, proceeding exactly as in section 3.2.ii, we see that

$$(2.5)\qquad y_1 - b_1 y_2 = B_1 e^{r_1 t} \quad ; \quad y_1 - b_2 y_2 = B_2 e^{r_2 t}$$

where

$$(2.6)\qquad b_1, b_2 = (1/2\nu_2)[(\sigma_2-\sigma_1) \pm \{(\sigma_1-\sigma_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}}] \quad ,$$

$$(2.7)\qquad r_1, r_2 = \tfrac{1}{2}[(\sigma_1+\sigma_2) \pm \{(\sigma_1-\sigma_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}}] \quad ,$$

B_1 and B_2 are constants of integration, $\sigma_i = \mu_i+\nu_i$ and $y_i = 1-z_i$ $(i = 1,2)$. Note that if $b_1 = b_2$ then $r_1 = r_2$ and expressions (2.5) are equivalent.

The last of the auxiliary equations (2.4) is given by

$$(2.8)\qquad dG/dt = -G[(\lambda_1 m_1(t)+\alpha_1)y_1 + (\lambda_2 m_2(t)+\alpha_2)y_2] \quad .$$

To solve this differential equation we first write expressions (2.3.11) and (2.3.12) in the form

$$(2.9)\qquad m_i(t) \equiv h_i e^{\omega_1 t} + k_i e^{\omega_2 t} + \ell_i \qquad (i = 1,2)$$

and replace (2.5) by the equivalent expressions

$$(2.10)\qquad \begin{aligned} y_1 &= (b_2-b_1)^{-1}(B_1 b_2 e^{r_1 t} - B_2 b_1 e^{r_2 t}) \\ y_2 &= (b_2-b_1)^{-1}(B_1 e^{r_1 t} - B_2 e^{r_2 t}) \quad . \end{aligned}$$

We now substitute for $m_i(t)$ and y_i $(i=1,2)$ from (2.9) and (2.10) into (2.8), and then integrate the resulting expression, thereby obtaining

$$
\begin{aligned}
(b_1-b_2)\log(G) + \text{const.} = B_1\{&(\lambda_1 h_1 b_2+\lambda_2 h_2)(\omega_1+r_1)^{-1}e^{(\omega_1+r_1)t} \\
&+ (\lambda_1 k_1 b_2+\lambda_2 k_2)(\omega_2+r_1)^{-1}e^{(\omega_2+r_1)t} \\
&+ [(\lambda_1 \ell_1+\alpha_1)b_2 + (\lambda_2 \ell_2+\alpha_2)]r_1^{-1}e^{r_1 t}\} \\
-B_2\{&(\lambda_1 h_1 b_1+\lambda_2 h_2)(\omega_1+r_2)^{-1}e^{(\omega_1+r_2)t} \\
&+ (\lambda_1 k_1 b_1+\lambda_2 k_2)(\omega_2+r_2)^{-1}e^{(\omega_2+r_2)t} \\
&+ [(\lambda_1 \ell_1+\alpha_1)b_1 + (\lambda_2 \ell_2+\alpha_2)]r_2^{-1}e^{r_2 t}\} \quad .
\end{aligned}
\tag{2.11}
$$

When $b_1 \neq b_2$ the two solutions (2.5) are different and may be combined with (2.11) to give the general solution

$$G(z_1,z_2;t) = f(B_1,B_2)\exp\{(b_1-b_2)^{-1}H(t)\} \tag{2.11a}$$

where f is an arbitrary function and $H(t)$ is given by the right hand side of (2.11), namely

$$(b_1-b_2)\log(G) + \text{const.} = H(t) \quad .$$

Expression (2.11a) is the general integral of the Lagrange linear equation (2.2), and to derive the required integral we find that solution (2.11a) which satisfies the initial condition (2.3). At time $t=0$ expressions (2.10) become

$$z_1 = 1 - (b_2-b_1)^{-1}(B_1 b_2 - B_2 b_1)$$

$$z_2 = 1 - (b_2-b_1)^{-1}(B_1 - B_2)$$

and so

$$f(B_1,B_2) = [1-(b_2-b_1)^{-1}(B_1b_2-B_2b_1)]^{a_1}[1-(b_2-b_1)^{-1}(B_1-B_2)]^{a_2}$$

$$\times \exp\{-(b_1-b_2)^{-1} H(0)\} \quad .$$

We now replace the constants B_1 and B_2 by expressions (2.5), namely

$$B_1 = (y_1-b_1y_2)e^{-r_1t} \quad ; \quad B_2 = (y_1-b_2y_2)e^{-r_2t} \quad ,$$

and obtain the following solution $(b_1 \neq b_2)$ for $G(z_1,z_2;t)$

$$\begin{aligned} G(z_1,z_2;t) = {} & [1-(b_2-b_1)^{-1}\{(1-z_1)(b_2e^{-r_1t}-b_1e^{-r_2t}) \\ & -(1-z_2)b_1b_2(e^{-r_1t}-e^{-r_2t})\}]^{a_1} \\ & \times [1-(b_2-b_1)^{-1}\{(1-z_1)(e^{-r_1t}-e^{-r_2t}) \\ & -(1-z_2)(b_1e^{-r_1t}-b_2e^{-r_2t})\}]^{a_2} \\ & \times \exp\{(b_1-b_2)^{-1}[(\omega_1+r_1)^{-1}(\lambda_1h_1b_2+\lambda_2h_2)\{(1-z_1)-b_1(1-z_2)\}(e^{\omega_1t}-e^{-r_1t}) \\ & + (\omega_2+r_1)^{-1}(\lambda_1k_1b_2+\lambda_2k_2)\{(1-z_1)-b_1(1-z_2)\}(e^{\omega_2t}-e^{-r_1t}) \\ & + r_1^{-1}[(\lambda_1\ell_1+\alpha_1)b_2+(\lambda_2\ell_2+\alpha_2)]\{(1-z_1)-b_1(1-z_2)\}(1-e^{-r_1t}) \\ & - (\omega_1+r_2)^{-1}(\lambda_1h_1b_1+\lambda_2h_2)\{(1-z_1)-b_2(1-z_2)\}(e^{\omega_1t}-e^{-r_2t}) \\ & -(\omega_2+r_2)^{-1}(\lambda_1k_1b_1+\lambda_2k_2)\{(1-z_1)-b_2(1-z_2)\}(e^{\omega_2t}-e^{-r_2t}) \\ & - r_2^{-1}[(\lambda_1\ell_1+\alpha_1)b_1+(\lambda_2\ell_2+\alpha_2)]\{(1-z_1)-b_2(1-z_2)\}(1-e^{-r_2t})]\} \quad . \end{aligned} \tag{2.12}$$

Note that $G(1,1;t) \equiv 1$ $(t \geq 0)$ as required. This solution has a relatively straightforward structure and moments of the process may be derived from it if both sides of (2.12) are differentiated with respect to the z_i the appropriate number of times. Moreover the probability $p_{ij}(t)$ may be evaluated as the coefficient of $z_1^i z_2^j$. In particular we see that if both colonies are initially empty, so that $a_1 = a_2 = 0$, the p.g.f. $G(z_1,z_2;t)$ factorizes as

$$G(z_1,z_2;t) = G_1(z_1;t)G_2(z_2;t)$$

where $G_1(z_1;t)$ and $G_2(z_2;t)$ are the p.g.f's of two independent Poisson processes.

4.2.ii Equilibrium approximations

For the modified process to possess an equilibrium solution it is clearly necessary and sufficient for the 'immigration' rates $\alpha_i+\lambda_i m_i(t)$, and hence the means $m_i(t)$, to be bounded above. This implies that $\omega_1 < 0$, and hence automatically $\omega_2 < 0$, whence it follows from (2.3.24), that for $\xi_i=\lambda_i-\mu_i-\nu_i$ $(i=1,2)$,

$$\begin{aligned} m_1(\infty) &= (\alpha_2\nu_2-\alpha_1\xi_2)/(\xi_1\xi_2-\nu_1\nu_2) = \ell_1 \\ m_2(\infty) &= (\alpha_1\nu_1-\alpha_2\xi_1)/(\xi_1\xi_2-\nu_1\nu_2) = \ell_2 \quad . \end{aligned} \tag{2.13}$$

Moreover it follows from (2.7) that both $r_1 > 0$ and $r_2 > 0$. Thus if $\omega_1 < 0$ then as $t \to \infty$ in (2.12) we obtain

$$G(z_1,z_2;\infty) = \exp\{(b_1-b_2)^{-1}\{r_1^{-1}[(\lambda_1\ell_1+\alpha_1)b_2+(\lambda_2\ell_2+\alpha_2)][(1-z_1)-b_1(1-z_2)]$$

$$- r_2^{-1}[(\lambda_1\ell_1+\alpha_1)b_1+(\lambda_2\ell_2+\alpha_2)][(1-z_1)-b_2(1-z_2)]\}\} ,$$

and this simplifies to give

$$G(z_1,z_2;\infty) = \exp\{\ell_1(z_1-1)+\ell_2(z_2-1)\} \tag{2.14}$$

where ℓ_1 and ℓ_2 are defined by (2.13). Thus in equilibrium the modified process behaves as two independent Poisson processes, the joint probabilities being given by

$$p_{ij}(\infty) = \frac{\ell_1^i}{i!}\,\frac{\ell_2^j}{j!}\,e^{-(\ell_1+\ell_2)} \quad . \tag{2.15}$$

4.2.iii The homogeneous model

We shall now use (2.12) to derive the probabilities $p_{ij}(t)$ for the spatially homogeneous model as they are easier to interpret than those for the general model. As we have already obtained the

general equilibrium solution (2.14), we shall suppress immigration ($\alpha_1=\alpha_2=0$) for the remaining part of this section in order to make our results more manageable.

Placing $\lambda_i=\lambda$, $\mu_i=\mu$, $\nu_i=\nu$, $\alpha_i=0$ $(i=1,2)$ in (2.6) and (2.7) we obtain

$$\begin{aligned} b_1&=1 \quad ; \quad b_2=-1 \\ r_1&=\mu+2\nu \quad ; \quad r_2=\mu \quad , \end{aligned} \tag{2.16}$$

whilst from (2.3.9) and (2.9) we get

$$\begin{aligned} \omega_1&=\lambda-\mu \quad ; \quad \omega_2=\lambda-\mu-2\nu \\ h_1&=h_2=\tfrac{1}{2}(a_1+a_2) \quad ; \quad k_1=-k_2=\tfrac{1}{2}(a_1-a_2) \quad . \end{aligned} \tag{2.17}$$

If we substitute from (2.16) and (2.17) into (2.12) we then get the solution

$$G(z_1,z_2;t)=[g_1(z_1,z_2;t)]^{a_1}[g_2(z_1,z_2;t)]^{a_2} \quad , \tag{2.18}$$

where for $i,j=1,2;\ j\neq i$

$$\begin{aligned} g_i(z_1,z_2;t)=&[1+\tfrac{1}{2}(z_i+z_j-2)e^{-\mu t}+\tfrac{1}{2}(z_i-z_j)e^{-(\mu+2\nu)t}] \\ &\times \exp\{\tfrac{1}{2}(e^{\lambda t}-1)[(z_i+z_j-2)e^{-\mu t}+(z_i-z_j)e^{-(\mu+2\nu)t}]\} \quad . \end{aligned} \tag{2.19}$$

This result was verified by back-substitution into the differential equation (2.2).

Thus, without loss of generality, we need only consider the initial conditions $a_1=1$, $a_2=0$. Writing

$$q=\tfrac{1}{2}e^{-\mu t}(1+e^{-2\nu t}) \quad ; \quad r=\tfrac{1}{2}e^{-\mu t}(1-e^{-2\nu t}) \quad ; \quad s=(e^{\lambda t}-1)$$

we have

$$g_1(z_1,z_2;t)=[(1-e^{-\mu t})+qz_1+rz_2]\exp\{s(qz_1+rz_2-e^{-\mu t})\} \quad , \tag{2.20}$$

and extracting the coefficient of $z_1^i z_2^j$ in $g_1(z_1,z_2;t)$ we get

$$(2.21)\quad p_{ij}(t) = q^i r^j s^{i+j-1}(i!j!)^{-1}[s(1-e^{-\mu t})+i+j]\exp\{-se^{-\mu t}\} \ .$$

We therefore see that for $\nu > 0$ and t large

$$p_{ij}(t) \sim (\tfrac{1}{2}e^{(\lambda-\mu)t})^{i+j}(i!j!)^{-1}\exp\{-e^{(\lambda-\mu)t}\} \ .$$

This is a bivariate Poisson distribution with common parameters $\tfrac{1}{2}\exp\{(\lambda-\mu)t\}$. Thus, for fixed $t > 0$, $p_{ij}(t)$ has its maximum value when i and j are both approximately equal to the integer part of $\tfrac{1}{2}\exp\{(\lambda-\mu)t\}$. Note that as $t \to \infty$ the migration rate ν has an exponentially small effect on the probabilities $p_{ij}(t)$.

We may evaluate the marginal probabilities $p_{i.}(t)$ and $p_{.j}(t)$ in the same way. For putting $z_2=1$ and $z_1=1$ respectively in (2.20) we get

$$g_1(z_1,1;t) = [(1-e^{-\mu t}+r)+qz_1]\exp\{s(qz_1+r-e^{-\mu t})\}$$

$$g_1(1,z_2;t) = [(1-e^{-\mu t}+q)+rz_2]\exp\{s(rz_2+q-e^{-\mu t})\} \ ,$$

and extracting the coefficient of z_1^i in $g_1(z_1,1;t)$ and z_2^j in $g_1(1,z_2;t)$ we have

$$(2.22)\qquad \begin{aligned} p_{i.}(t) &= q^i s^{i-1}(i!)^{-1}[s(1-e^{-\mu t}+r)+i]\exp\{s(r-e^{-\mu t})\} \\ p_{.j}(t) &= r^j s^{j-1}(j!)^{-1}[s(1-e^{-\mu t}+q)+j]\exp\{s(q-e^{-\mu t})\} \ . \end{aligned}$$

Still assuming that $\alpha_1=\alpha_2=0$ together with the initial conditions $a_1=1$ and $a_2=0$, let us define $p_{00}(t)$ and $\bar{p}_{00}(t)$ to be the probabilities, associated with equation (2.2.2) and equation (2.2) respectively, that both colonies are empty at time $t \geq 0$. As $\lambda_1=\lambda_2$ and $\mu_1=\mu_2$ the probability of extinction $p_{00}(t)$ is clearly identical with that of a single isolated colony. Thus

$$(2.23)\qquad p_{00}(t) = (\mu-\mu e^{-(\lambda-\mu)t})(\lambda-\mu e^{-(\lambda-\mu)t})^{-1}$$

(Cox and Miller (1965), p. 166), whilst from (2.21) we have

(2.24) $$\bar{p}_{00}(t) = (1-e^{-\mu t})\exp\{-e^{-\mu t}(e^{\lambda t}-1)\} .$$

The simple two-colony birth-death-migration process and the modified process have one major difference between them. Birth rates in the former are proportional to the sizes of the two colonies and so once both colonies are empty the population becomes extinct. However, birth rates in the latter are proportional to the two means $m_i(t)$ $(i=1,2)$ and are therefore independent of the actual colony sizes. In this sense birth is equivalent to time-dependent immigration. Extinction is clearly impossible, as

$$\Pr[\text{a birth in the time interval } (t,\infty)] > 0$$

for all $t \geq 0$.

For $\lambda > \mu$, $m_1(t)$ and $m_2(t)$ increase exponentially for $t \geq t_0$ (some $t_0 > 0$), and so $\bar{p}_{00}(t)$ will quickly tend to zero. For $\lambda < \mu$, $m_1(t)$ and $m_2(t)$ decrease exponentially for $t \geq t_1$ (some $t_1 > 0$) , and so we expect that $\bar{p}_{00}(t)$ will quickly tend to one. If $\lambda=\mu$, $m_i(t) \sim \frac{1}{2}$ and the process reduces to a simple immigration-death process for which an equilibrium distribution should develop. Thus, although the probability of ultimate extinction is zero, we expect $\bar{p}_{00}(t) \to \bar{p}$ for some $0 < \bar{p} < 1$. In fact, from (2.23) and (2.24) we have

$$\begin{aligned}
&\bar{p}_{00}(\infty) = 0 \quad , \quad p_{00}(\infty) = \mu/\lambda \qquad (\lambda > \mu) \\
&\bar{p}_{00}(\infty) = e^{-1} \quad , \quad p_{00}(\infty) = 1 \qquad (\lambda = \mu) \\
&\bar{p}_{00}(\infty) = 1 \quad , \quad p_{00}(\infty) = 1 \qquad (\lambda < \mu) \quad .
\end{aligned}$$

4.2.iv Moments

To obtain moments of the modified process either we may differentiate the p.g.f. (2.12) with respect to z_1 and z_2 the

required number of times and then place $z_1=z_2=1$, or we may work with the original differential equation (2.2). Now the coefficients of z_1 and z_2 in expression (2.12) do not possess a simple form, and even the demonstration that

$$\left.\frac{\partial G(z_1,z_2;t)}{\partial z_1}\right/_{z_1=z_2=1} = m_1(t)$$

when $\alpha_1=\alpha_2=0$ involves a great amount of tedious algebra (not shown here). Moreover, the work associated with the derivation of moments of general order, when $\alpha_1=\alpha_2\neq 0$, would be considerably greater. Let us, therefore, return to equation (2.2).

As it is usually easier to work with the cumulant generating function (c.g.f.) rather than to work with the probability generating function, let us replace the dummy variables z_1 and z_2 in equation (2.2) by e^{θ_1} and e^{θ_2} respectively. The equation for the c.g.f.

$$K(\theta_1,\theta_2;t) = \log[G(z_1,z_2;t)]$$

is then given by

$$\frac{\partial K}{\partial t} = \frac{\partial K}{\partial \theta_1}[\mu_1(e^{-\theta_1}-1)+\nu_1(e^{\theta_2-\theta_1}-1)] + \frac{\partial K}{\partial \theta_2}[\mu_2(e^{-\theta_2}-1)+\nu_2(e^{\theta_1-\theta_2}-1)]$$

$$+[(\lambda_1 m_1(t)+\alpha_1)(e^{\theta_1}-1) + (\lambda_2 m_2(t)+\alpha_2)(e^{\theta_2}-1)] \quad . \tag{2.25}$$

To obtain the first-order moments we differentiate (2.25) with respect to θ_i $(i=1,2)$ and place $\theta_1=\theta_2=0$. This procedure yields the two differential equations

$$d\bar{m}_i(t)/dt = \lambda_i m_i(t) - (\mu_i+\nu_i)\bar{m}_i(t) + \nu_j \bar{m}_j(t) + \alpha_i \quad (i=1,2;\ j\neq i) \ , \tag{2.26}$$

where $m_i(t)$ and $\bar{m}_i(t)$ denote the means of the ordinary and modified process respectively. A comparison of equations (2.3.1) and (2.26) gives

$$\bar{m}_i(t) = m_i(t) \quad (i=1,2; \quad t \geq 0) \quad ,$$

and so the approximation leaves the means unaltered. This result is easily shown by letting

$$M_i(t) = \bar{m}_i(t) - m_i(t) \quad (i=1,2) \quad .$$

For subtracting (2.3.1) from (2.26) we obtain

$$\begin{aligned} dM_1(t)/dt &= -(\mu_1+\nu_1)M_1(t) + \nu_2 M_2(t) \\ dM_2(t)/dt &= -(\mu_2+\nu_2)M_2(t) + \nu_1 M_1(t) \end{aligned} \tag{2.27}$$

with

$$M_1(0) = M_2(0) = 0 \quad .$$

Writing equations (2.27) in the vector-matrix form

$$d\underline{M}(t)/dt = A\underline{M}(t)$$

and integrating, where $\underline{M}(t) = (M_1(t), M_2(t))$ and A is a constant matrix, we have

$$\underline{M}(t) = \underline{M}(0)\exp\{At\} = \underline{0} \quad ,$$

the null vector.

This equivalence does not hold for the second-order moments, however. For differentiating (2.25) a second time with respect to θ_1 and θ_2 and placing $\theta_1=\theta_2=0$ we get, for $\alpha_1=\alpha_2=0$,

$$\begin{aligned} d\bar{v}_{11}(t)/dt &= (\lambda_1+\mu_1+\nu_1)m_1(t) + \nu_2 m_2(t) - 2(\mu_1+\nu_1)\bar{v}_{11}(t) + 2\nu_2\bar{v}_{12}(t) \\ d\bar{v}_{12}(t)/dt &= -\nu_1 m_1(t) - \nu_2 m_2(t) + \nu_1\bar{v}_{11}(t) - (\mu_1+\mu_2+\nu_1+\nu_2)\bar{v}_{12}(t) \\ &\qquad + \nu_2\bar{v}_{22}(t) \\ d\bar{v}_{22}(t)/dt &= (\lambda_2+\mu_2+\nu_2)m_2(t) + \nu_1 m_1(t) - 2(\mu_2+\nu_2)\bar{v}_{22}(t) + 2\nu_1\bar{v}_{12}(t) \quad , \end{aligned} \tag{2.28}$$

where the $\bar{v}_{ij}(t)$ represent the variances and covariance of the modified process. Note that the only difference between equations (2.28) and (2.4.3) - (2.4.5) is that the coefficients of the $\bar{v}_{ij}(t)$ have $\lambda_1=\lambda_2=0$. Applying the Laplace transformation (2.3.3) to

equations (2.28) we have

$$(2.29)\quad s\bar{v}^*_{11}(s) = (\lambda_1+\mu_1+\nu_1)m^*_1(s)+\nu_2 m^*_2(s)-2(\mu_1+\nu_1)\bar{v}^*_{11}(s)+2\nu_2\bar{v}^*_{12}(s)$$

$$(2.30)\quad s\bar{v}^*_{12}(s) = -\nu_1 m^*_1(s) - \nu_2 m^*_2(s)+\nu_1\bar{v}^*_{11}(s)-(\mu_1+\nu_1+\mu_2+\nu_2)\bar{v}^*_{12}(s)$$

$$+\nu_2\bar{v}^*_{22}(s)$$

$$(2.31)\quad s\bar{v}^*_{22}(s) = (\lambda_2+\mu_2+\nu_2)m^*_2(s)+\nu_1 m^*_1(s)-2(\mu_2+\nu_2)\bar{v}^*_{22}(s)+2\nu_1\bar{v}^*_{12}(s)$$

for $\Re(s)$ large enough to ensure the existence of $m^*_i(s)$ and $\bar{v}^*_{ij}(s)$ $(i,j=1,2)$. Writing equations (2.3.4) and (2.3.5), with $\alpha_1=\alpha_2=0$, in the form

$$(2.32)\quad \begin{aligned} sm^*_1(s) - a_1 &= (\lambda_1-\mu_1-\nu_1)m^*_1(s) + \nu_2 m^*_2(s) \\ sm^*_2(s) - a_2 &= (\lambda_2-\mu_2-\nu_2)m^*_2(s) + \nu_1 m^*_1(s) \quad , \end{aligned}$$

and combining (2.32) with (2.29) and (2.31), we get

$$(2.33)\quad \begin{aligned} \bar{v}^*_{11}(s)[\,s+2(\mu_1+\nu_1)] &= 2\nu_2\bar{v}^*_{12}(s) + [s+2(\mu_1+\nu_1)]\,m^*_1(s)-a_1 \\ \bar{v}^*_{22}(s)[\,s+2(\mu_2+\nu_2)] &= 2\nu_1\bar{v}^*_{12}(s) + [s+2(\mu_2+\nu_2)]\,m^*_2(s)-a_2 \quad . \end{aligned}$$

Substituting for $\bar{v}^*_{11}(s)$ and $\bar{v}^*_{22}(s)$ from (2.33) into (2.30) and writing

$$\sigma_i = \mu_i + \nu_i \qquad (i=1,2)$$

we obtain

$$(2.34)\quad \bar{v}^*_{12}(s)\,[s+(\sigma_1+\sigma_2)]\,[(s+2\sigma_1)(s+2\sigma_2)-4\nu_1\nu_2]$$

$$= -\nu_1 a_1(s+2\sigma_2) - \nu_2 a_2(s+2\sigma_1) \quad .$$

But the equation

$$s^2 + 2s(\sigma_1+\sigma_2) + 4(\sigma_1\sigma_2-\nu_1\nu_2) = 0$$

has roots

$$(2.35)\quad s_1, s_2 = -(\sigma_1+\sigma_2) \pm \{(\sigma_1-\sigma_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}} \quad .$$

Thus (2.34) may be written as

$$\bar{V}^{*}_{12}(s) = [-\nu_1 a_1(s+2\sigma_2) - \nu_2 a_2(s+2\sigma_1)]/(s+\sigma_1+\sigma_2)(s-s_1)(s-s_2) \ .$$

This expression is easily inverted with the inverse transformation

$$(s-a)^{-1} \equiv \mathcal{L}[e^{at}]$$

after expansion into partial fractions. If $s_1 \neq s_2$, this procedure gives the solution

$$\begin{aligned} \bar{V}_{12}(t) = \{ & 2e^{-(\sigma_1+\sigma_2)t}(\sigma_2-\sigma_1)(\nu_1 a_1 - \nu_2 a_2) \\ & - e^{s_1 t}[\nu_1 a_1(s_1+2\sigma_2) + \nu_2 a_2(s_1+2\sigma_1)] \\ & - e^{s_2 t}[\nu_1 a_1(s_2+2\sigma_2) + \nu_2 a_2(s_2+2\sigma_1)]\}/2[(\sigma_1-\sigma_2)^2 + 4\nu_1\nu_2] \ . \end{aligned} \tag{2.36}$$

Provided $s_1 \neq s_2$ it follows from (2.35) that

$$s_1 > -(\sigma_1+\sigma_2) > s_2 \ ,$$

whilst

$$s_1 = -(\sigma_1+\sigma_2) + \{(\sigma_1+\sigma_2)^2 - 4(\mu_1\mu_2+\nu_1\mu_2+\nu_2\mu_1)\}^{\frac{1}{2}} < 0 \ . \tag{2.37}$$

Hence, when t is large

$$\bar{V}_{12}(t) \sim -e^{s_1 t}[\nu_1 a_1(s_1+2\sigma_2) + \nu_2 a_2(s_1+2\sigma_1)]/2[(\sigma_1-\sigma_2)^2 + 4\nu_1\nu_2] \ , \tag{2.38}$$

and so the covariance

$$\bar{V}_{12}(t) \to 0 \quad \text{as} \quad t \to \infty \ .$$

To evaluate the variances $\bar{V}_{11}(t)$ and $\bar{V}_{22}(t)$ we write the differential equations (2.28) in the integral form

$$\bar{V}_{11}(t) = \int_0^t e^{2\sigma_1(s-t)}[(\lambda_1+\mu_1+\nu_1)m_1(s) + \nu_2 m_2(s) + 2\nu_2 \bar{V}_{12}(s)]\,ds \tag{2.39}$$

$$\bar{V}_{22}(t) = \int_0^t e^{2\sigma_2(s-t)}[(\lambda_2+\mu_2+\nu_2)m_2(s) + \nu_1 m_1(s) + 2\nu_1 \bar{V}_{12}(s)]\,ds \tag{2.40}$$

and replace $m_1(t)$, $m_2(t)$ and $\bar{V}_{12}(t)$ by expressions (2.3.15), (2.3.16) and (2.36), respectively. The asymptotic behaviour of $\bar{V}_{11}(t)$ and $\bar{V}_{22}(t)$ is clearly determined by the dominant exponents in the integrands of expressions (2.39) and (2.40). Now from (2.3.9)

$$\omega_1 = \tfrac{1}{2}[(\xi_1+\xi_2) + \{(\xi_1-\xi_2)^2 + 4\nu_1\nu_2\}^{\frac{1}{2}}]$$

and so

$$\partial\omega_1/\partial\lambda_1 = \tfrac{1}{2}[1 + \{(\xi_1-\xi_2)^2 + 4\nu_1\nu_2\}^{-\frac{1}{2}}(\xi_1-\xi_2)] > 0$$

$$\partial\omega_1/\partial\lambda_2 = \tfrac{1}{2}[1 + \{(\xi_1-\xi_2)^2 + 4\nu_1\nu_2\}^{-\frac{1}{2}}(\xi_2-\xi_1)] > 0 \ .$$

Hence the minimum value of ω_1 for fixed σ_1 and σ_2 is at $\lambda_1=\lambda_2=0$. This is to be expected as the minimum value of ω_1 corresponds to the minimum rate of growth of the mean population sizes. Thus

$$\begin{aligned}(\omega_1-s_1) &\geq \tfrac{1}{2}[-(\sigma_1+\sigma_2) + \{(\sigma_1-\sigma_2)^2 + 4\nu_1\nu_2\}^{\frac{1}{2}}] - s_1 \\ &= -\tfrac{1}{2}s_1 \qquad \text{(from (2.35))} \\ &> 0 \qquad \text{(from (2.37))} \ . \end{aligned} \tag{2.41}$$

The dominant exponents in the integrands of expressions (2.39) and (2.40) are therefore

$$\exp\{s(2\sigma_1+\omega_1)\} \quad \text{and} \quad \exp\{s(2\sigma_2+\omega_1)\} \ ,$$

respectively. Substituting for $m_1(t)$ and $m_2(t)$ from (2.9), with $\ell_1=\ell_2=0$, into (2.39) we obtain

$$\bar{V}_{11}(t) \sim e^{-2\sigma_1 t} \int^t e^{(2\sigma_1+\omega_1)s} [(\lambda_1+\mu_1+\nu_1)h_1 + \nu_2 h_2]\,ds \ ,$$

with a similar expression holding for $\bar{V}_{22}(t)$. These expressions integrate directly to give for $s_1 \neq s_2$

$$\begin{aligned}\bar{V}_{11}(t) &\sim e^{\omega_1 t}\,[(\lambda_1+\mu_1+\nu_1)h_1 + \nu_2 h_2]/[\omega_1 + 2(\mu_1+\nu_1)] \\ \bar{V}_{22}(t) &\sim e^{\omega_1 t}\,[(\lambda_2+\mu_2+\nu_2)h_2 + \nu_1 h_1]/[\omega_1 + 2(\mu_2+\nu_2)] \ . \end{aligned} \tag{2.42}$$

Combining expressions (2.36) and (2.42) we see that the linear correlation coefficient

$$\bar{\rho}(t) = \bar{V}_{12}(t)/\{\bar{V}_{11}(t)\bar{V}_{22}(t)\}^{\frac{1}{2}}$$

is of order

$$\bar{\rho}(t) = O(e^{(s_1-\omega_1)t}) \ .$$

It therefore follows from (2.41) that as $t \to \infty$

$$\bar{\rho}(t) \to 0$$

irrespective of whether the total mean population size is asymptotically bounded or unbounded. Also, substituting for $m_1(t)$ and $m_2(t)$ from (2.9) with $\ell_1=\ell_2=0$ into (2.42) we have

$$\bar{v}_{11}(t)/m_1(t) \sim [(\lambda_1+\mu_1+\nu_1) + \nu_2(h_2/h_1)]/[\omega_1 + 2(\mu_1+\nu_1)]$$

$$\bar{v}_{22}(t)/m_2(t) \sim [(\lambda_2+\mu_2+\nu_2) + \nu_1(h_1/h_2)]/[\omega_1 + 2(\mu_2+\nu_2)] \quad ,$$

and so the ratio of the variance to the mean population size in each colony is asymptotically constant.

Although the modification to the birth process leaves the means unaltered, it clearly has a considerable effect on the second-order moments. For it follows from (2.4.17) and (2.4.18) that if $\omega_1 > 0$ the variances and covariance of the unmodified process are all of equal order, namely

$$v_{ij}(t) = O(e^{2\omega_1 t}) \qquad (i,j=1,2) \quad ,$$

whilst from (2.38) and (2.42)

$$\bar{v}_{11}(t), \bar{v}_{22}(t) = O(e^{\omega_1 t}) \; ; \quad \bar{v}_{12}(t) = O(e^{s_1 t}) \quad .$$

Thus the order of the variances of the modified process is the square-root of that of the unmodified process. Moreover, not only is the covariance $\bar{v}_{12}(t)$ now of smaller order than the variances, but

$$\bar{v}_{12}(t) \to 0 \quad \text{as} \quad t \to \infty$$

irrespective of the sign of ω_1 .

For the spatially homogeneous model with the parameter values $\lambda_i=\lambda$, $\mu_i=\mu$ and $\nu_i=\nu$ $(i=1,2)$, the second-order moments simplify to give

$$\bar{V}_{11}(t) = e^{(\lambda-\mu)t}(h+ke^{-2\nu t}) - \tfrac{1}{2}e^{-2\mu t}(h+2ke^{-2\nu t}+he^{-4\nu t})$$

(2.43) $$\bar{V}_{12}(t) = -\tfrac{1}{2}he^{-2\mu t}(1-e^{-4\nu t})$$

$$\bar{V}_{22}(t) = e^{(\lambda-\mu)t}(h-ke^{-2\nu t}) - \tfrac{1}{2}e^{-2\mu t}(h-2ke^{-2\nu t}+he^{-4\nu t})$$

where $h=\frac{1}{2}(a_1+a_2)$ and $k=\frac{1}{2}(a_1-a_2)$. These expressions clearly differ considerably from the variances and covariance, $V_{ij}(t)$, defined by expression (2.5.25). Note that the underlying condition $s_1 \neq s_2$ is equivalent to $\nu > 0$, and when this holds the correlation coefficient $\bar{\rho}(t)$ is given by

$$\bar{\rho}(t) \sim -1 \qquad \text{if } \lambda=\mu=0$$

$$\bar{\rho}(t) \sim -\tfrac{1}{2}e^{-(\lambda+\mu)t} \qquad \text{otherwise .}$$

The first of these results is intuitively reasonable. For if $\lambda=\mu=0$ the colonies develop as a simple bivariate immigration-migration process, and so the total population size is monotonic increasing and unbounded. Thus as t becomes large the

$$\Pr[\text{a particular event is a migration}] \to 1 ,$$

whence a decrease in one colony necessarily results in an increase in the other. If $\nu=0$ it follows from the second of equations (2.28) that $\bar{V}_{12}(t) \equiv 0$, as would be expected.

Expression (2.36) and all subsequent results depend on the assumption that $s_1 \neq s_2$. If $s_1 = s_2$, it follows from (2.35) that both $\mu_1+\nu_1=\mu_2+\nu_2$ and $\nu_1\nu_2=0$. Migration may therefore occur in one direction only, so that $\nu_2=0$ (say), and equations (2.28) may each be integrated in turn. The analysis parallels that given in section 2.4.iv for the unmodified process.

4.3 The Negative Binomial Approximation

In the previous section we modified the birth mechanism and this led to a distribution which was related to the bivariate Poisson

distribution. Although the means of the original and modified processes are identical their second-order moments are different. The question then arises whether a different form of approximating probability generating function can be found which leaves both the first- and second-order moments intact. Bartlett (1949) suggests fitting a negative binomial distribution for the one-dimensional multiplicative process, and we shall now determine whether or not this type of distribution yields a reasonable approximation to our two-colony model.

Consider the p.g.f.

$$(3.1) \qquad G(z_1,z_2;t) = \left(\frac{1-b_1z_1}{1-b_1}\right)^{d_1}\left(\frac{1-b_2z_2}{1-b_2}\right)^{d_2} ,$$

with associated probabilities

$$(3.2) \qquad p_{ij}(t) = (1-b_1)^{-d_1}(1-b_2)^{-d_2}\, b_1^i b_2^j \binom{i-d_1-1}{-d_1-1}\binom{j-d_2-1}{-d_2-1} ,$$

where

$$(3.3) \qquad 0 < b_i < 1 \quad \text{and} \quad d_i < 0 \qquad (i=1,2)$$

are time-dependent parameters. Expression (3.1) represents the product of the p.g.f's of two independent negative binomial processes, and so cannot be an exact solution as the colony sizes are interdependent. However, under certain circumstances it may well yield a good approximation to the solution of the Kolmogorov forward equation (2.2.2).

Let us try to choose the parameters b_i and d_i $(i=1,2)$ so that the first- and second-order moments derived from (3.1) are the same as the $m_i(t)$ and $V_{ij}(t)$. Differentiating (3.1) with respect to z_1 and z_2 and placing $z_1=z_2=1$ in the usual manner we get the

relations

$$m_1(t) = -b_1 d_1/(1-b_1) \tag{3.4}$$

$$m_2(t) = -b_2 d_2/(1-b_2) \tag{3.5}$$

with

$$V_{11}(t) - m_1(t) + m_1^2(t) = b_1^2 d_1(d_1-1)/(1-b_1)^2 \tag{3.6}$$

$$V_{12}(t) + m_1(t)m_2(t) = b_1 b_2 d_1 d_2/(1-b_1)(1-b_2) \tag{3.7}$$

$$V_{22}(t) - m_2(t) + m_2^2(t) = b_2^2 d_2(d_2-1)/(1-b_2)^2 \quad . \tag{3.8}$$

On substituting for $m_1(t)$ and $m_2(t)$ from (3.4) and (3.5) we have

$$V_{12}(t) \equiv 0 \ , \tag{3.9}$$

as would be expected, together with

$$V_{11}(t) - m_1(t) + m_1^2(t) = m_1^2(t)(d_1-1)/d_1$$

$$V_{22}(t) - m_2(t) + m_2^2(t) = m_2^2(t)(d_2-1)/d_2$$

which invert to give

$$\begin{aligned} d_1 &= m_1^2(t)/(m_1(t)-V_{11}(t)) \\ d_2 &= m_2^2(t)/(m_2(t)-V_{22}(t)) \ . \end{aligned} \tag{3.10}$$

Equations (3.4) and (3.5) may now be solved for b_1 and b_2, yielding

$$\begin{aligned} b_1 &= 1 - m_1(t)/V_{11}(t) \\ b_2 &= 1 - m_2(t)/V_{22}(t) \ . \end{aligned} \tag{3.11}$$

Combining (3.10) and (3.11) with (3.1) we get

$$G(z_1,z_2;t) = \prod_{i=1}^{2} \{m_i^{-1}[V_{ii}+(m_i-V_{ii})z_i]\}^{m_i^2/(m_i-V_{ii})} \tag{3.12}$$

where we have omitted the variable t for typographical convenience.

The representation (3.1) for the approximating p.g.f. contains only 4 parameters and yet we are using 5 moment equations to estimate them. This imposes a constraint on the moments, namely that the

covariance between the two colonies is zero. This will be approximately true when the migration rates are small in comparison with the birth and death rates so that the colonies are 'slightly connected' (see section 5.2).

If there exist specific values of b_i and d_i $(i=1,2)$, determined from (3.10) and (3.11), which do not satisfy conditions (3.3), then expression (3.1) will not have the required negative binomial form. An investigation into the existence of such values would clearly be useful. However, we have previously shown in Chapter 2 that if the dominant eigen-value ω_1 is positive then

$$m_i(t) = O(e^{\omega_1 t}) > 0 \quad ; \quad V_{ii}(t) = O(e^{2\omega_1 t}) > 0 \qquad (i=1,2;\ t>0) .$$

Thus, as $t \to \infty$

$$d_i \sim -m_i^2(t)/V_{ii}(t)$$

asymptotically approaches a negative constant, whilst b_i converges to 1 from below. Hence, conditions (3.3) are certainly satisfied if $\omega_1 > 0$ provided that t is reasonably large. If ω_1 is negative then an equilibrium distribution exists and it follows, from (3.10) and (3.11), that in the equilibrium situation we require $m_i(\infty) < V_{ii}(\infty)$ $(i=1,2)$. This seems intuitively reasonable, especially in the light of several simulation runs (for an example, see section 2.2.ii).

The equilibrium p.g.f. of a simple birth-death-immigration process, with parameters λ, μ and α respectively, is given by the expression

$$[(\mu-\lambda z)/(\mu-\lambda)]^{-\alpha/\lambda} \tag{3.13}$$

(see, for example, Bartlett (1966), p. 82). Thus as $t \to \infty$ expression (3.12) represents the product of the p.g.f's of any two independent simple birth-death-immigration processes whose parameters λ, μ and α

satisfy the relations

$$m_i(\infty) = \alpha/(\mu-\lambda) \quad \text{and} \quad V_{ii}(\infty) = \alpha\mu/(\mu-\lambda)^2 \qquad (i=1,2) \quad .$$

Moreover, if migration is suppressed expression (3.12) becomes the exact solution as $t\to\infty$. For placing $\nu_1=\nu_2=0$ in expressions (2.3.24) and (2.4.21)-(2.4.23) we get for $\omega_1 < 0$

$$m_i(\infty) = -\alpha_i/(\lambda_i-\mu_i) \qquad (i=1,2)$$

$$V_{ii}(\infty) = \alpha_i\mu_i/(\lambda_i-\mu_i)^2 \qquad (i=1,2)$$

$$V_{12}(\infty) = 0 \ ,$$

and on letting $t\to\infty$ in (3.12) we have

$$(3.14) \qquad G(z_1,z_2;\infty) = \left[\frac{\mu_1-\lambda_1 z_1}{\mu_1-\lambda_1}\right]^{-\alpha_1/\lambda_1} \left[\frac{\mu_2-\lambda_2 z_2}{\mu_2-\lambda_2}\right]^{-\alpha_2/\lambda_2} \quad .$$

In order to fit the covariance term $V_{12}(t)$ we require an approximate representation for $G(z_1,z_2;t)$ which contains more parameters than (3.1), for example

$$G(z_1,z_2;t) = \left[\frac{1-b_1z_1-c_1z_2}{1-b_1-c_1}\right]^{d_1} \left[\frac{1-b_2z_1-c_2z_2}{1-b_2-c_2}\right]^{d_2} \quad .$$

This expression does not yield a readily solvable set of equations for the unknown constants b_i, c_i and d_i $(i=1,2)$, and further study of more suitable representations would be desirable.

The negative binomial approximation possesses two distinct advantages over the modified process discussed in section 4.2. In the first place the probabilities (3.2) have a very simple algebraic structure, whilst the probabilities derived from the coefficients of $z_1^i z_2^j$ in expression (2.12) are relatively complex. Secondly, the two variances associated with the approximation (3.12) are exact,

whilst those of the modified process are not even of the right order of magnitude if $\omega_1 > 0$. However, the modified process does at least have a physical interpretation, and this must surely count in its favour. I shall make a more detailed comparison between the two approximations in section 7.3.

4.4 Approximations based on Bounded Population Size

For any $\varepsilon > 0$ and $0 < T < \infty$, we may choose a positive integer $N(\varepsilon,T)$ such that, for $t \in [0,T]$,

$$\sum_{i+j>N} p_{ij}(t) < \varepsilon \quad . \tag{4.1}$$

Thus approximations to the probabilities $p_{ij}(t)$ may, in principle, be obtained to any desired degree of accuracy by introducing a sufficiently large upper bound N on the total population size, and solving the resulting finite set of difference-differential equations.

4.4.i The equilibrium approximation

Assume an equilibrium situation exists with the associated limiting probabilities

$$\pi_{ij} = \lim_{t\to\infty} p_{ij}(t) \qquad (0 \leq i,j < \infty) \quad .$$

Then letting $t \to \infty$ in equations (2.2.1) we obtain the set of equations

$$\begin{aligned} \pi_{ij}[i(\lambda_1+\mu_1+\nu_1)+j(\lambda_2+\mu_2+\nu_2)+\alpha_1+\alpha_2] &= \lambda_1(i-1)\pi_{i-1,j} + \lambda_2(j-1)\pi_{i,j-1} \\ &+ \mu_1(i+1)\pi_{i+1,j} + \mu_2(j+1)\pi_{i,j+1} + \nu_1(i+1)\pi_{i+1,j-1} \\ &+ \nu_2(j+1)\pi_{i-1,j+1} + \alpha_1\pi_{i-1,j} + \alpha_2\pi_{i,j-1} \end{aligned} \tag{4.2}$$

which are valid for $0 \leq i,j < \infty$.

If we now introduce a maximum population size N , by suppressing

birth and immigration whenever $i+j=N$, equations (4.2) will still hold provided that $0 \leq i+j < N$. However, if $i+j=N$ we have to replace them by the set

$$\begin{aligned} \pi_{ij}[i(\mu_1+\nu_1)+j(\mu_2+\nu_2)] &= \lambda_1(i-1)\pi_{i-1,j} + \lambda_2(j-1)\pi_{i,j-1} \\ &+ \nu_1(i+1)\pi_{i+1,j-1} + \nu_2(j+1)\pi_{i-1,j+1} \qquad (i+j=N) \\ &+ \alpha_1\pi_{i-1,j} + \alpha_2\pi_{i,j-1} \quad . \end{aligned} \tag{4.3}$$

Terms involving λ_i and α_i $(i=1,2)$ are omitted from the left hand side of expression (4.3) as the total population size $i+j$ is not permitted to increase from N to $N+1$. Similarly, terms involving μ_1 and μ_2 are omitted from the right hand side as it is not possible for $i+j$ to decrease from $N+1$ to N.

Equations (4.2) and (4.3), together with the normalising condition

$$\sum_{i=0}^{N} \sum_{j=0}^{N-i} \pi_{ij} = 1 \quad , \tag{4.4}$$

describe a finite irreducible Markov chain and hence possess a unique solution. Numerical approximations to the original unbounded system of equations, represented by equations (4.2) with $0 \leq i,j < \infty$, may therefore be obtained by writing equations (4.2)-(4.4) in vector-matrix form and then using standard numerical procedures. However, in Chapter 7 we show how to determine the π_{ij} by using simulation techniques, and this method would appear to be by far the better way to obtain particular numerical solutions.

Suppose that the death rates μ_i are large compared with the birth and immigration rates λ_i and α_i. Reasonable approximations to the π_{ij} may then be obtained when the upper bound N is sufficiently small for the solution of equations (4.2)-(4.4) to

be algebraically feasible. In particular, suppose that a spatially homogeneous population is divided into two colonies, so that for $i=1,2$ we have $\lambda_i=\lambda$, $\mu_i=\mu$, $\nu_i=\nu$ and $\alpha_i=\alpha$. Now the equilibrium p.g.f. of a simple birth-death-immigration process, with parameters λ, μ and 2α respectively, is given by

$$G(z) = \left(\frac{1-\rho z}{1-\rho}\right)^{-2\alpha/\lambda} \tag{4.5}$$

where $\rho=\lambda/\mu$ (see expression (3.13)). Thus, the equilibrium probabilities

$$\pi_n = \Pr[\text{total combined population is of size } n] \quad (n=0,1,2\ldots)$$

are given by

$$\pi_n = (1-\lambda/\mu)^{2\alpha/\lambda}(\mu^n n!)^{-1}(2\alpha)(2\alpha+\lambda)\ldots(2\alpha+(n-1)\lambda) \; . \tag{4.6}$$

In these circumstances migration clearly has no effect on the total population size so that

$$\pi_n = \sum_{i=0}^{n} \pi_{i,n-i} \qquad (n=0,\ldots,N) \; , \tag{4.7}$$

whilst from symmetry

$$\pi_{ij} = \pi_{ji} \; . \tag{4.8}$$

Result (4.7) was verified by placing $j=n-i$ in expressions (4.2) and (4.3), summing both sides over $i=0,\ldots,n$, and then using an induction argument to derive (4.6) in the form

$$\pi_n = \pi_0(\mu^n n!)^{-1}(2\alpha)\ldots(2\alpha+(n-1)\lambda) \qquad (n=0,\ldots,N) \; .$$

On using (4.7) and (4.8) we may easily solve equations (4.2) and (4.3) to give, for example,

$N \geq 2$: $\quad \pi_{01} = \pi_{10} = (\alpha/\mu)\pi_{00}$

$N = 2$: $\quad \pi_{20} = \pi_{02} = (\alpha/2\mu^2)[\alpha+\lambda(\mu+\nu)/(\mu+2\nu)]\pi_{00}$

(4.9) $\pi_{11} = (\alpha/\mu^2)[\alpha+\lambda\nu/(\mu+2\nu)]\pi_{00}$

N = 3 : $\pi_{20} = \pi_{02} = (\alpha/2\mu^2)[(\lambda+\alpha)(\lambda+2\alpha)(\mu+2\nu) + (\lambda+2\alpha)(\mu+\nu)(3\mu+4\nu)$

$- \alpha\mu(3\mu+4\nu)]\pi_{00}/[(3\mu+4\nu)(\lambda+\mu+2\nu+\alpha)-\mu(2\lambda+\alpha)]$

$\pi_{11} = -2\pi_{20} + (\alpha/\mu^2)(2\alpha+\lambda)\pi_{00}$.

We may either let π_{00} retain its true value (4.6), namely

(4.10) $$\pi_{00} = (1-\lambda/\mu)^{2\alpha/\lambda} ,$$

or replace it by the value determined from condition (4.4). For example, if N=2 the equation

$$\sum_{n=0}^{2} \sum_{i=0}^{2-n} \pi_{i,n-i} = 1$$

yields

(4.11) $$\pi_{00} = \mu^2/[\mu^2+2\alpha\mu+\alpha(\lambda+2\alpha)] .$$

Note that (4.10) expands to give

$$\pi_{00}^{-1} = 1 + 2\alpha\mu^{-1} + \alpha(\lambda+2\alpha)\mu^{-2} + O(\mu^{-3}) ,$$

which agrees with (4.11) to order $O(\mu^{-3})$.

4.4.ii <u>The general approximation</u>

This procedure of approximate solutions for bounded total population size may be generalized to cover the probabilities $p_{ij}(t)$. For let the generating function of the colony sizes at time $t \geq 0$, conditional on the total population being of size n , be

(4.12) $$g_n(z;t) = \sum_{i=0}^{n} p_{n-i,i}(t) z^i \quad (n=0,1,2,\ldots)$$

and denote $\sigma_i = \lambda_i + \mu_i + \nu_i$ $(i=1,2)$. Then from equations (2.2.1) we may obtain the differential-difference equation

$$\frac{\partial g_n}{\partial t} = [\lambda_1(n-1)+\alpha_1+\alpha_2 z]g_{n-1} + z(\lambda_2 z-\lambda_1)\frac{\partial g_{n-1}}{\partial z} + [n\nu_1 z-(n\sigma_1+\alpha_1+\alpha_2)]g_n$$

(4.13)
$$+ [\nu_2+(\sigma_1-\sigma_2)z-\nu_1 z^2]\frac{\partial g_n}{\partial z} + \mu_1(n+1)g_{n+1} + (\mu_2-\mu_1 z)\frac{\partial g_{n+1}}{\partial z} .$$

Integrating equation (4.13) with respect to z we get the iterative solution

$$g_{n+1}(z;t) = (\mu_2-\mu_1 z)^{n+1}c_{n+1}(t) + (\mu_2-\mu_1 z)^{n+1}\int^z (\mu_2-\mu_1 y)^{-(n+2)}\left\{\frac{\partial g_n(y;t)}{\partial t}\right.$$

(4.14)
$$+ (n\sigma_1+\alpha_1+\alpha_2-n\nu_1 y)g_n(y;t) + [\nu_1 y^2 + (\sigma_2-\sigma_1)y-\nu_2]\frac{\partial g_n(y;t)}{\partial y}$$

$$\left. - [(n-1)\lambda_1+\alpha_1+\alpha_2 y]g_{n-1}(y;t) + y(\lambda_1-\lambda_2 y)\frac{\partial g_{n-1}(y;t)}{\partial y}\right\} dy$$

where the $c_n(t)$ are the 'constants' of integration.

The problem now revolves around the determination of the unknown set $\{c_n(t)\}$. In fact the probabilities $p_{0,n+1}(t)$ may be regarded as the remaining unknowns. For writing the integral in (4.14) as

$$\int^z (\mu_2-\mu_1 y)^{-(n+2)}Q_n(y;t)dy ,$$

we see that

$$Q_n(y;t) = y^{n+1}(-n\nu_1 p_{0n}(t) + \nu_1 n p_{0n}(t)) + \text{poly.(degree } n \text{ in } y)$$

$$= \text{poly.(degree } n \text{ in } y) .$$

Thus for some functions $\{b_{in}(t)\}$

$$g_{n+1}(z;t) = (\mu_2-\mu_1 z)^{n+1}[c_{n+1}(t) + \int^z (\sum_{j=2}^{n+2} b_{jn}(t)(\mu_2-\mu_1 y)^{-j})dy]$$

(4.15)
$$= c_{n+1}(t)(\mu_2-\mu_1 z)^{n+1} + \text{poly.(degree } n \text{ in } z) ,$$

and comparing (4.12) with (4.14) we have

$$p_{0,n+1}(t) = c_{n+1}(t)(-\mu_1)^{n+1} \quad . \tag{4.16}$$

We now place an upper bound N on the maximum total population size, where N satisfies condition (4.1). For $0 \leq n < N$ the difference-differential equations (2.2.1), and hence expressions (4.14), remain unchanged. For $n=N$ they become modified to

$$\begin{aligned} dp_{N-i,i}(t)/dt = {} & [\lambda_1(N-i-1)+\alpha_1]p_{N-i-1,i}(t) + [\lambda_2(i-1)+\alpha_2]p_{N-i,i-1}(t) \\ & + \nu_1(N-i+1)p_{N-i+1,i-1}(t) + \nu_2(i+1)p_{N-i-1,i+1}(t) \\ & - [(N-i)(\mu_1+\nu_1) + i(\mu_2+\nu_2)]p_{N-i,i}(t) \quad , \end{aligned} \tag{4.17}$$

and on multiplying both sides by z^i and summing over the range $i=0,\ldots,N$ we see that equations (4.17) may be written in the form

$$\begin{aligned} & N[\mu_1+\nu_1(1-z)]g_N(z;t) + [(\mu_2-\mu_1)z+\nu_2(z-1)+\nu_1 z(z-1)]\frac{\partial g_N(z;t)}{\partial z} + \frac{\partial g_N(z;t)}{\partial t} \\ & = [\lambda_1(N-1)+\alpha_1+\alpha_2 z]g_{N-1}(z;t) - z(\lambda_1-\lambda_2 z)\frac{\partial g_{N-1}(z;t)}{\partial z} \quad . \end{aligned} \tag{4.18}$$

Denote $\{g_n^{(N)}(z;t)\}$ for $n=0,1,\ldots,N$ as the sequence of generating functions generated recursively from equations (4.14) conditional on the total population size being bounded above by N. Then $\{g_n^{(N)}(z;t)\}$ is a sequence of polynomials of degree n in z, linear in the unknown functions $c_n(t)$ and their derivatives. The $c_n(t)$ $(0 \leq n \leq N)$ may be determined, at least in principle, if we substitute for $g_N^{(N)}(z;t)$ and $g_{N-1}^{(N)}(z;t)$ in equation (4.18) and then equate like coefficients of z^i $(0 \leq i \leq N)$ to zero. Finally we use the normalizing condition

$$\sum_{n=0}^{N} g_n^{(N)}(1;t) \equiv 1 \qquad (t \geq 0) \quad .$$

Note that in the above analysis we write $g_{-1}(z;t) \equiv 0$ together

with $g_0(z;t) = p_{00}(t) \equiv c_0(t)$ $(t \geq 0)$. The unrestricted solution $(N=\infty)$ is formally given by the limit

$$g_n(z;t) = \lim_{N\to\infty} g_n^{(N)}(z;t) \ .$$

4.4.iii Illustration of the general approximation with N=1

For the purpose of illustration we shall place an upper bound of N=1 on the total population size and assume that both colonies are initially empty. Thus the only possible states are 00, 01 and 10. Substituting for

$$p_{00}(t) \equiv c_0(t) \tag{4.19}$$

in expression (4.14) with $n=0$ we get

$$g_1(z;t) = (\mu_2-\mu_1 z)c_1(t) + \mu_1^{-1}[c_0'(t)+(\alpha_1+\alpha_2)c_0(t)] \tag{4.20}$$

(here a dash (') denotes differentiation with respect to t). Placing (4.19) and (4.20) in equation (4.18) with $N=1$ and equating coefficients of z^i $(i=0,1,2)$ we obtain the two differential equations

$$(\nu_1\mu_2+\mu_1\mu_2+\mu_1\nu_2)c_1(t) + \mu_1 c_1'(t) + [(\nu_1/\mu_1)(\alpha_1+\alpha_2)+\alpha_2]c_0(t) + (\nu_1/\mu_1)c_0'(t) = 0 \tag{4.21}$$

$$(\nu_1\mu_2+\mu_1\mu_2+\mu_1\nu_2)c_1(t) + \mu_2 c_1'(t) + [(1+\nu_1/\mu_1)(\alpha_1+\alpha_2)-\alpha_1]c_0(t) + (1/\mu_1)[(\mu_1+\nu_1)+(\alpha_1+\alpha_2)]c_0'(t) + (1/\mu_1)c_0''(t) = 0 \ . \tag{4.22}$$

When $\mu_1=\mu_2=\mu$ (say) equations (4.21) and (4.22) have identical terms in $c_1(t)$ and $c_1'(t)$ and hence are easily integrated by standard methods to give the following solution:

$$p_{00}(t) \equiv c_0(t) = 1 - (\alpha_1+\alpha_2)(1-e^{-(\mu+\alpha_1+\alpha_2)t})/(\mu+\alpha_1+\alpha_2) \ ,$$

$$p_{01}(t) = -\mu c_1(t) \ , \tag{4.23}$$

$$p_{10}(t) = \mu c_1(t) + (1/\mu)[c_0'(t)+(\alpha_1+\alpha_2)c_0(t)]$$

where

$$c_1(t) = (\mu+\alpha_1+\alpha_2)^{-1}(\mu+\nu_1+\nu_2)^{-1}[(\nu_1/\mu)(\alpha_1+\alpha_2)+\alpha_2](e^{-(\mu+\nu_1+\nu_2)t}-1)$$

$$+ (\alpha_1+\alpha_2)(\nu_1-\alpha_2)(e^{-(\mu+\alpha_1+\alpha_2)t} - e^{-(\mu+\nu_1+\nu_2)t})/\mu(\mu+\alpha_1+\alpha_2)(\nu_1+\nu_2-\alpha_1-\alpha_2) .$$

Expressions (4.23) were verified by back-substitution into the difference-differential equations (2.2.1) and (4.17) for the $p_{n-i,i}(t)$.

Except in simple cases, like the previous example, the above technique of bounding the population size will not produce solutions for the probabilities $p_{ij}(t)$; for the determination of the unknown functions $\{c_n(t)\}$ will in general be totally impracticable if not impossible. However, we have already seen in the introduction to section 3.5 that provided the matrix of transition rates has the appropriate finite triangular structure Severo's technique will provide the solution. Thus, provided the matrix of transition rates can be transformed into the required form, approximate solutions for the $p_{ij}(t)$ can be developed if we first place an upper bound on the total population size and then apply Severo's approach.

CHAPTER 5

OTHER MODES OF APPROXIMATING THE SOLUTIONS TO THE KOLMOGOROV FORWARD EQUATIONS : TWO-COLONY MODELS

5.1 Introduction

The Kolmogorov forward and backward differential equations for the probabilities $p_{ij}(t)$ are intractable in the sense that we are unable to write down an explicit expression for their general solution. However, we may obtain solutions of a recursive nature and in this chapter I consider two such approaches. First I shall express the solutions to the backward equations as power series expansions of the common migration rate ν (this result is dependent on $\nu_1=\nu_2=\nu$) . Here the coefficients of ν^n $(n=0,1,2,\ldots)$ are defined recursively by means of two integral equations. Secondly I shall give a brief description of the derivation of recursive sequences of functions which converge to the required solution.

5.2 Migration Models with Slightly Connected Colonies

5.2.i Solutions to the Kolmogorov backward equations

So far we have been working with the Kolmogorov forward differential equation (2.2.2). By considering the first event which occurs in the period $(0,t)$, Puri (1968) obtains the corresponding backward equations

$$(2.1)\quad dg_i(t)/dt = \lambda_i g_i^2(t) - (\lambda_i+\mu_i+\nu_i)g_i(t) + \nu_i g_j(t) + \mu_i \qquad (i=1,2;j\neq i)$$

for the special case in which immigration is suppressed $(\alpha_1=\alpha_2=0)$.

Here the functions $g_i(z_1,z_2;t)$ $(i=1,2)$ denote the probability generating functions of the probabilities $p_{ij}(t)$, subject to the initial conditions $X_i(0)=1$, $X_j(0)=0$ $(j\neq i)$.

Suppose that the migration rates ν_i are equal. Then on putting $\nu_1=\nu_2=\nu$ we may write equations (2.1) in the form

$$\begin{aligned} dg_1(t)/dt &= \lambda_1 g_1^2(t) - (\lambda_1+\mu_1)g_1(t) + \mu_1 + \nu[g_2(t)-g_1(t)] \\ dg_2(t)/dt &= \lambda_2 g_2^2(t) - (\lambda_2+\mu_2)g_2(t) + \mu_2 + \nu[g_1(t)-g_2(t)] \end{aligned} \qquad (2.2)$$

If $\nu=0$ it follows that equations (2.2) represent two independent birth-death processes and hence have known solutions. The introduction of a migration rate ν which is small in comparison with the λ_i and μ_i should cause only a slight perturbation of these known p.g.f's. We shall therefore express $g_1(t)$ and $g_2(t)$ as power series in ν, for when ν is small it should be necessary to derive the first few terms only.

An example of such 'slightly connected' colonies is represented by two populations which exist in two neighbouring valleys separated only by a high mountain ridge. Migration from one valley to the other is possible but rare.

Denote

$$g_1(t) = \sum_{i=0}^{\infty} a_i(t)\nu^i \quad \text{and} \quad g_2(t) = \sum_{i=0}^{\infty} b_i(t)\nu^i \,, \qquad (2.3)$$

where $a_i(t)$ and $b_i(t)$ are functions of z_1, z_2 and t but are independent of the parameter ν. Replacing $g_1(t)$ and $g_2(t)$ in equations (2.2) by their power series expansions (2.3), and equating like coefficients of ν^r $(r=0,1,2,\ldots)$ on both sides of the

resulting expressions we obtain the following set of differential-difference equations

$$(2.4)\quad da_0/dt = \lambda_1 a_0^2 - (\lambda_1+\mu_1)a_0 + \mu_1$$

$$(2.5)\quad da_n/dt = a_n[\,2a_0\lambda_1 - (\lambda_1+\mu_1)] + \lambda_1 \sum_{r=1}^{n-1} a_r a_{n-r} + b_{n-1} - a_{n-1} \qquad (n>0)\ ,$$

where $\sum_{r=1}^{n-1} a_r a_{n-r} \equiv 0$ for $n=1$. Because of the obvious similarity between $a_n(t)$ and $b_n(t)$ most results are stated for $a_n(t)$ alone. At $t=0$ we have $g_i(z_1,z_2;0)=z_i$ $(i=1,2)$, and so the initial conditions are given by

$$(2.6)\quad \begin{aligned} a_0(0) &= z_1\ ,\ b_0(0) = z_2 \\ a_i(0) &= b_i(0) \qquad (i>0) \end{aligned}\ .$$

We see from (2.4) that $a_0(t)$ and $b_0(t)$ denote the p.g.f's of two unconnected simple birth-death processes $(\nu=0)$. Hence it immediately follows that

$$(2.7)\quad a_0(t) = \frac{\mu_1(1-z_1)-(\mu_1-\lambda_1 z_1)\exp\{-(\lambda_1-\mu_1)t\}}{\lambda_1(1-z_1)-(\mu_1-\lambda_1 z_1)\exp\{-(\lambda_1-\mu_1)t\}}$$

(Cox and Miller (1965), p.165).

Now the solution of the standard differential equation

$$dx(t)/dt + P(t)x(t) = Q(t) \quad : \quad x(0) = 0$$

is given by

$$x(t) = \int_0^t Q(s)\exp\{-\int_s^t P(u)du\}ds$$

(see, for example, Ince (1956)). Hence on denoting

$$P(u) = -[\,2\lambda_1 a_0(u) - (\lambda_1+\mu_1)]$$

$$Q(s) = \lambda_1 \sum_{r=1}^{n-1} a_r(s)a_{n-r}(s) + b_{n-1}(s) - a_{n-1}(s)\ ,$$

we see that equation (2.5) integrates to give

$$(2.8)\quad a_n(t) = \int_0^t [b_{n-1}(s)-a_{n-1}(s)+\lambda_1 \sum_{r=1}^{n-1} a_r(s)a_{n-r}(s)] \times \exp\{\int_s^t [2\lambda_1 a_0(u)-(\lambda_1+\mu_1)]\,du\}ds$$

for $n=1,2,\ldots,$ with a comparable expression for $b_n(t)$. The solution of the differential equations (2.2) is, therefore, given by the power series expansions (2.3) where the coefficients $a_n(t)$ and $b_n(t)$ are evaluated recursively from (2.7) and (2.8).

In particular, if we put $n=1$ we obtain the first term of the 'perturbation' from the single isolated colony situation, namely

$$(2.9)\quad a_1(t) = \int_0^t [b_0(s)-a_0(s)] \exp\{\int_s^t [2\lambda_1 a_0(u)-(\lambda_1+\mu_1)]\,du\}ds \quad .$$

In order to evaluate this integral let us first denote

$$(2.10)\quad c_i = \mu_i-\lambda_i \quad , \quad d_i = \lambda_i(1-z_i) \quad , \quad h_i = \mu_i-\lambda_i z_i \quad , \quad r_i = \mu_i/\lambda_i$$

where $i=1,2$, and define the integral

$$(2.11)\qquad R(x,y;a,t) = \int_0^t e^{xs}(1-ae^{ys})^{-1}ds \quad .$$

We begin by evaluating the function

$$(2.12)\qquad I(s,t) = \int_s^t [2\lambda_1 a_0(u)-(\lambda_1+\mu_1)]\,du$$

which is common to all of the integrals (2.8). Substituting for $a_0(u)$ from (2.7) into (2.12) and integrating we get

$$(2.13)\qquad \exp\{I(s,t)\} = e^{c_1(t-s)}(d_1-h_1e^{c_1 s})^2(d_1-h_1e^{c_1 t})^{-2} \quad .$$

Thus

$$J_0(t) \equiv \int_0^t a_0(s)\exp\{I(s,t)\}ds$$

$$= e^{c_1 t}(d_1-h_1e^{c_1 t})^{-2}[-r_1d_1^2c_1^{-1}(e^{-c_1 t}-1)-(r_1+1)d_1h_1t$$

$$+h_1^2c_1^{-1}(e^{c_1 t}-1)] \quad . \tag{2.14}$$

Similarly

$$\int_0^t b_0(s)\exp\{I(s,t)\}ds = e^{c_1 t}(d_1-h_1e^{c_1 t})^{-2}[\int_0^t e^{-c_1 s}(d_1-h_1e^{c_1 s})^2ds$$

$$+ d_2(r_2-1)\int_0^t e^{-c_1 s}(d_1-h_1e^{c_1 s})^2(d_2-h_2e^{c_2 s})^{-1}ds] \tag{2.15}$$

$$\equiv J_1(t) + I_2(t) \qquad \text{(say)} \quad .$$

Integrating expression (2.15) and denoting

$$I_1(t) \equiv J_1(t) - J_0(t)$$

we obtain the solution

$$g_1(z_1,z_2;t) = a_0(t) + \nu[I_1(t)+I_2(t)] + O(\nu^2) \quad , \tag{2.16}$$

where $a_0(t)$ is given by (2.7),

$$I_1(t) = e^{c_1 t}(1-z_1)(d_1-h_1e^{c_1 t})^{-2}[d_1(e^{-c_1 t}-1)+h_1c_1t] \tag{2.17}$$

and

$$I_2(t) = e^{c_1 t}(r_2-1)(d_1-h_1e^{c_1 t})^{-2}[d_1^2R(-c_1,c_2;h_2/d_2,t)$$

$$+ h_1^2R(c_1,c_2;h_2/d_2,t) - 2(d_1h_1/c_2)\log\{e^{c_2 t}(d_2-h_2)/(d_2-h_2e^{c_2 t})\}] \quad . \tag{2.18}$$

The dummy variables z_1 and z_2 are suitably chosen to ensure the existence of expressions (2.17) and (2.18).

Note that if we put $z_1=z_2=1$ in (2.17) and (2.15) we get

$I_1(t) \equiv 0$ and $I_2(t) \equiv 0$ for $t \geq 0$ and so, to order ν,

$$g_1(1,1;t) \equiv 1 \tag{2.19}$$

as required.

The integral (2.11) is related to the Incomplete Beta function and so in general can not be evaluated in closed form. However, for the particular case $\mu_1-\lambda_1=\mu_2-\lambda_2=c$ (say)

$$I_2(t) = e^{ct}(r_2-1)(cd_2)^{-1}(d_1-h_1e^{ct})^{-2}[ct(d_1^2h_2-2d_1d_2h_1) + d_1^2d_2(1-e^{-ct}) - h_2^{-1}(d_2h_1-d_1h_2)^2\log\{(d_2-h_2e^{ct})/(d_2-h_2)\}] , \tag{2.20}$$

and so (2.16) may be evaluated exactly for given values of the parameters.

The above results on 'slightly connected processes' have previously been derived for the special case of the spatially homogeneous model in which $\lambda_1=\lambda_2=\lambda$ and $\mu_1=\mu_2=\mu$ as well as $\nu_1=\nu_2=\nu$. The first-order solution corresponding to (2.16) was determined, in which $I_2(t)$ has an even simpler form than that given in expression (2.20), and bounds were developed for the corresponding second-order terms $a_2(t)$ and $b_2(t)$. (See Renshaw (1970, 1973a)).

5.2.ii Probability of extinction

The individual probabilities $p_{ij}(t)$ may be retrieved from expression (2.16) for the p.g.f. $g_1(z_1,z_2;t)$ by evaluating the coefficients of $z_1^iz_2^j$. In particular, the probability of extinction $p_{00}(t)$ is obtained by placing $z_i=0$, $d_i=\lambda_i$ and $h_i=\mu_i$ $(i=1,2)$ in (2.16). This procedure gives

(2.21) $$p_{00}(t) = a_0(t) + \nu[I_1(t)+I_2(t)] + O(\nu^2) \quad ,$$

where we now have

$$a_0(t) = \mu_1(1-e^{c_1 t})/(\lambda_1-\mu_1 e^{c_1 t})$$

together with

$$I_1(t) = e^{c_1 t}(\lambda_1-\mu_1 e^{c_1 t})^{-2}[\lambda_1(e^{-c_1 t}-1)+\mu_1 c_1 t]$$

and

$$I_2(t) = e^{c_1 t}(r_2-1)(\lambda_1-\mu_1 e^{c_1 t})^{-2}[\lambda_1^2 R(-c_1,c_2;\mu_2/\lambda_2,t) + \mu_1^2 R(c_1,c_2;\mu_2/\lambda_2,t) - 2(\lambda_1\mu_1/c_2)\log\{e^{c_2 t}(\lambda_2-\mu_2)/(\lambda_2-\mu_2 e^{c_2 t})\}] \quad .$$

5.2.iii Marginal distributions for the homogeneous model

Let us consider the effect of the small migration rate $\nu > 0$ on the marginal distributions of the spatially homogeneous model for which $\lambda_1=\lambda_2=\lambda$ and $\mu_1=\mu_2=\mu$. On placing $z_2=1$ we see that expressions (2.10) become

$$d_1 = \lambda(1-z_1) \quad , \quad h_1 = \mu-\lambda z_1 \quad , \quad c_1 = \mu-\lambda$$
$$d_2 = 0 \quad , \quad h_2 = \mu-\lambda \quad , \quad c_2 = \mu-\lambda \quad ,$$

whilst from (2.15)

$$I_2(t) \equiv 0 \qquad (t \geq 0) \quad .$$

Expression (2.16) then yields the marginal p.g.f.

(2.22) $$g_1(z_1,1;t) = a_0(t) + \nu(1-z_1)[\lambda(1-z_1)(1-e^{(\mu-\lambda)t}) + (\mu-\lambda z_1)(\mu-\lambda)te^{(\mu-\lambda)t}]/[\lambda(1-z_1)-(\mu-\lambda z_1)e^{(\mu-\lambda)t}]^2 + O(\nu^2) \quad .$$

Similarly, to evaluate the marginal p.g.f. $g_1(1,z_2;t)$ we place $z_1=1$ in (2.10), whence

$$d_1 = 0 \quad , \quad h_1 = \mu-\lambda \quad , \quad c_1 = \mu-\lambda$$
$$d_2 = \lambda(1-z_2) \quad , \quad h_2 = \mu-\lambda z_2 \quad , \quad c_2 = \mu-\lambda \quad .$$

Clearly from (2.7) and (2.17)

$$a_0(t) \equiv 1 \quad \text{and} \quad I_1(t) \equiv 0 \qquad (t \geq 0) \quad .$$

Expressions (2.16) and (2.20) then yield

$$(2.23)\quad g_1(1,z_2;t) = 1 - \nu e^{-(\mu-\lambda)t}(1-z_2)(\mu-\lambda z_2)^{-1}\log[\{\lambda(1-z_2) - (\mu-\lambda z_2)e^{(\mu-\lambda)t}\}/(\lambda-\mu)] + O(\nu^2) \quad .$$

Let us recall that associated with the p.g.f. $g_1(z_1,z_2;t)$ are the initial conditions

$$X_1(0) = 1 \quad ; \quad X_2(0) = 0 \quad .$$

Thus if colony 1 becomes empty in the unconnected case ($\nu=0$) the population automatically becomes extinct. However, if $\nu > 0$ this is not necessarily so. For provided colony 2 is not empty, an individual may migrate from it to colony 1 and so restart the population there. Denote the probability that colony 1 is empty at time $t \geq 0$ by

$$p_{0.}(t) = \Pr[X_1(t)=0 \mid \nu=0]$$
$$\bar{p}_{0.}(t) = \Pr[X_1(t)=0 \mid \nu>0] \quad .$$

Then putting $z_1=0$ in expression (2.22) we obtain the standard result

$$(2.24)\qquad p_{0.}(t) = (\mu-\mu e^{(\mu-\lambda)t})/(\lambda-\mu e^{(\mu-\lambda)t})$$

(Cox and Miller (1965), p.166) together with

$$(2.25)\quad \bar{p}_{0.}(t) = p_{0.}(t) + \nu[\lambda+e^{(\mu-\lambda)t}\{\mu(\mu-\lambda)t-\lambda\}]/(\lambda-\mu e^{(\mu-\lambda)t})^2 + O(\nu^2) \quad .$$

It follows that to order ν

$$\bar{p}_{0.}(t) > p_{0.}(t) \qquad (t > 0) \quad .$$

Moreover, if $\mu > \lambda$ we have to order ν ,

$$\bar{p}_{0.}(\infty) = p_{0.}(\infty) = 1 \quad ,$$

whilst if $\mu < \lambda$

$$p_{0.}(\infty) = \mu/\lambda$$
$$\bar{p}_{0.}(\infty) = (\mu+\nu)/\lambda + O(\nu^2) \quad ,$$

an increase of approximately ν/λ .

Similarly, denoting the probability that colony 2 is empty at time $t \geq 0$ by

$$p_{.0}(t) = \Pr[X_2(t)=0 \mid \nu=0]$$
$$\bar{p}_{.0}(t) = \Pr[X_2(t)=0 \mid \nu>0]$$

and placing $z_2=0$ in expression (2.23), we have

$$p_{.0}(t) \equiv 1 \qquad (t \geq 0) \quad ,$$

as would be expected, together with

(2.26) $$\bar{p}_{.0}(t) = 1 - (\nu/\mu)e^{(\lambda-\mu)t}\log[(\lambda-\mu e^{(\mu-\lambda)t})/(\lambda-\mu)] + O(\nu^2) \quad .$$

Results (2.25) and (2.26) were also obtained by first placing $z_1=0$, $z_2=1$ and $z_1=1$, $z_2=0$, respectively, in equation (2.5) with $n=1$, and then integrating the resulting expressions.

Although for certain values of the parameters λ_i, μ_i and ν it will be sufficient to consider the first two terms alone in the expansion of $g_1(z_1,z_2;t)$ (namely $a_0(t)$ and $a_1(t)$) for other values it will be necessary to derive higher-order terms using the relation (2.8). As an illustration we see from (2.22) and (2.23)

that whilst both $a_1(z_1,1;t)$ and $a_1(1,z_2;t)$ remain bounded when $\mu > \lambda$, if $\mu < \lambda$ then the function $|a_1(1,z_2;t)|$ goes to infinity exponentially fast as t tends to infinity. Thus in this situation second- and higher-order terms cannot be neglected unless ν is extremely small.

That some breakdown like this should occur is not unreasonable. For although a change in the migration rate from $\nu=0$ to $0 < \nu << \lambda,\mu$ clearly results in only a minor perturbation of the population size of colony 1, the effect on colony 2 is considerable. Remembering that the initial conditions are $X_1(0)=1$ and $X_2(0)=0$, we have

$$X_2(t) \equiv 0 \tag{2.27}$$

for all $t \geq 0$ if $\nu=0$. If $\nu > 0$, but $\lambda < \mu$, the whole population becomes extinct with probability one, and so

$$\Pr[X_2(\infty)=0] = 1 \quad .$$

However, when $\nu > 0$ and $\lambda > \mu$, the mean size of colony 1 increases exponentially and the probability of a migration from colony 1 to colony 2 will eventually become appreciable no matter how small ν is. Once a migration occurs the mean size of colony 2 will then also increase exponentially, in direct contrast to (2.27).

Note that if $|z_1| < 1$, expressions (2.17) and (2.20) yield

$$I_1(t) + I_2(t) \to 0$$

as $t \to \infty$, regardless of whether $\mu > \lambda$ or $\mu < \lambda$. Thus $a_1(z_1,z_2;t)$ remains bounded provided that $|z_1| < 1$.

5.2.iv Admissible migration rates

The above results raise the question, "Over what values of ν and t is the first-order approximation (2.16) reasonable?". An answer to this question may be found from an examination of the first-order moments of the slightly connected process, and these may be obtained in the usual way by differentiating the function $g_i(z_1,z_2;t)$ given by expression (2.16) with respect to z_1 and z_2, and then placing $z_1=z_2=1$. In particular, let us consider the spatially homogeneous model with $\lambda_1=\lambda_2=\lambda$ and $\mu_1=\mu_2=\mu$. Here it is clearly sufficient to differentiate $g_1(z_1,z_2;t)$. If we denote the resulting mean population sizes of colonies 1 and 2 by $\bar{m}_1(t)$ and $\bar{m}_2(t)$, this procedure gives

$$\begin{aligned} \bar{m}_1(t) &= (1-\nu t)e^{(\lambda-\mu)t} + O(\nu^2) \\ \bar{m}_2(t) &= \nu t e^{(\lambda-\mu)t} + O(\nu^2) \quad . \end{aligned} \tag{2.28}$$

Placing $a_1=1$, $a_2=0$ and $\nu_1=\nu_2=\nu$ in (2.3.34) we obtain the corresponding exact expressions for the mean population sizes, namely

$$\begin{aligned} m_1(t) &= \tfrac{1}{2}e^{(\lambda-\mu)t}(1+e^{-2\nu t}) \\ m_2(t) &= \tfrac{1}{2}e^{(\lambda-\mu)t}(1-e^{-2\nu t}) \quad . \end{aligned} \tag{2.29}$$

These expand in powers of ν to give

$$\begin{aligned} m_1(t) &= e^{(\lambda-\mu)t}(1-\nu t) + O(\nu^2) \\ m_2(t) &= e^{(\lambda-\mu)t}(\nu t) + O(\nu^2) \quad . \end{aligned} \tag{2.30}$$

A comparison of (2.28) and (2.30) shows that $\bar{m}_i(t)=m_i(t)$ $(i=1,2)$ to order ν. The general result that

$$m_i(t) = \left[\frac{d}{dz_i}\sum_{r=0}^{j} a_r(t)\nu^r\right]\Bigg/_{z_1=z_2=1} + O(\nu^{j+1})$$

for all $j=0,1,2,\ldots$ follows directly from (2.3).

Let us base our answer to this question on the ratio defined by the function

$$\varrho_i^{(j)}(\nu t) = m_i(t)/m_i^{(j)}(t) \qquad (i=1,2) \quad , \tag{2.31}$$

where $m_i^{(j)}(t)$ represents the expansion of $m_i(t)$ up to and including terms of order ν^j. The proximity of the $\varrho_i^{(1)}(\nu t)$ to 1 indicates the success of the first-order approximation (2.16). Any improvement in the approximation by the inclusion of second-order terms may be determined from a comparison of $\varrho_i^{(2)}(\nu t)$ and $\varrho_i^{(1)}(\nu t)$ for $i=1,2$. It follows from (2.28) and (2.29) that for the spatially homogeneous model

$$\begin{aligned} \varrho_1^{(1)}(\nu t) &= \tfrac{1}{2}(1+e^{-2\nu t})/(1-\nu t) \\ \varrho_2^{(1)}(\nu t) &= (1-e^{-2\nu t})/(2\nu t) \end{aligned} \tag{2.32}$$

and

$$\begin{aligned} \varrho_1^{(2)}(\nu t) &= \tfrac{1}{2}(1+e^{-2\nu t})/[1-\nu t+(\nu t)^2] \\ \varrho_2^{(2)}(\nu t) &= \tfrac{1}{2}(1-e^{-2\nu t})/[\nu t-(\nu t)^2] \quad . \end{aligned} \tag{2.33}$$

The values of these expressions are given in table 5.1 for various values of νt.

Suppose we use the particular criterion

$$0.95 \leq \varrho_i^{(j)}(\nu t) \leq 1.05 \qquad (i=1,2)$$

to determine those values of νt for which the expansion

Table 5.1

Values of the function $\varrho_i^{(j)}(\nu t)$

(for explanation see text)

νt	$\varrho_1^{(1)}(\nu t)$	$\varrho_2^{(1)}(\nu t)$	$\varrho_1^{(2)}(\nu t)$	$\varrho_2^{(2)}(\nu t)$
0	1	1	1	1
0.01	1.000	0.990	1.000	1.000
0.05	1.003	0.952	1.000	1.002
0.10	1.010	0.907	0.999	1.007
0.15	1.024	0.864	0.998	1.017
0.20	1.044	0.824	0.994	1.030
0.25	1.071	0.787	0.989	1.049
0.50	1.368	0.632	0.912	1.264
0.75	2.446	0.518	0.753	2.072
1.00	∞	0.432	0.567	∞

$\Sigma_{i=0}^{j} a_i(t)\nu^i$ is a good approximation to $g_1(z_1,z_2;t)$. We have

$\varrho_1^{(1)}(0.21) = 1.049$ and $\varrho_2^{(1)}(0.05) = 0.952$ whence $0 \leq \nu t \leq 0.05$,

$\varrho_1^{(2)}(0.41) = 0.950$ and $\varrho_2^{(2)}(0.25) = 1.049$ whence $0 \leq \nu t \leq 0.25$.

These numerical results support our predictions of the previous sub-section that changes in νt affect colony 2 more than colony 1. They also show that for this particular model inclusion of the second-order term $a_2(t)$ results in a five-fold increase in the 'admissible range' of νt .

5.3 Sequences of Integral Equations

In the previous section we developed a power series solution to the backward equations (2.1) in the form

$$(2.3) \qquad g_1(t) = \sum_{n=0}^{\infty} a_n(t)\nu^n \quad ; \quad g_2(t) = \sum_{n=0}^{\infty} b_n(t)\nu^n$$

where the functions $a_n(t)$ and $b_n(t)$ are generated recursively from expression (2.8). Attempts to find other such solutions, for example an expansion in powers of t, have not proved successful. However, solutions of a different kind have been obtained by the use of convergent sequences of integral equations (see Puri (1968) and Renshaw (1970, 1973a)). We shall conclude this chapter with a brief description of their work.

Let the vector $\underline{X}(t)=(X_1(t),X_2(t))$ represent the sizes of the two colonies at time $t\geq 0$, and write $\sigma_i=\lambda_i+\mu_i+\nu_i$ $(i=1,2)$. Then, using the terminology of Feller (1966), denote

$$Q_t^{(n)}(\underline{x},\underline{w}) = \Pr[\text{of a transition from } \underline{X}(0)=\underline{x} \text{ to } \underline{X}(t)=\underline{w} \text{ in at most } n \text{ jumps}],$$

and write the p.g.f.

$$g_{it}^{(n)}(\underline{z}) = \sum_{w_1=0}^{\infty}\sum_{w_2=0}^{\infty} Q_t^{(n)}(\underline{e}_i,\underline{w})z_1^{w_1}z_2^{w_2},$$

where the vectors are defined as $\underline{e}_1=(1,0)$, $\underline{e}_2=(0,1)$, $\underline{w}=(w_1,w_2)$ and $\underline{z}=(z_1,z_2)$. The sequence of p.g.f's $\{g_{it}^{(n)}(\underline{z})\}$, defined by the recurrence relation

$$(3.1)\quad g_{it}^{(n+1)} = z_i e^{-\sigma_i t} + \int_0^t e^{-\sigma_i(t-s)}[\lambda_i(g_{is}^{(n)})^2 + \mu_i + \nu_i g_{js}^{(n)}]ds$$

$(i,j=1,2;\ j\neq i)$ with

$$(3.2)\quad g_{it}^{(0)} = z_i e^{-\sigma_i t},$$

is monotonic increasing in n and converges to the solution of the backward equations (2.1) (Renshaw (1970, 1973a)).

The same integral equation (3.1) was also obtained by the use of a technique similar to that used by Puri (1968). This time the

corresponding sequence of p.g.f's $\{f_{it}^{(n)}(\underline{z})\}$ converges monotonically downwards to the required solution, and condition (3.2) is replaced by

$$f_{it}^{(0)} = 1 \quad . \tag{3.3}$$

The technique is to subdivide each colony i into a countably infinite number of sub-colonies (states) S_{in} $(i=1,2;\ n=0,1,2,\ldots)$ and to consider the first event that happens to an initial individual in state S_{in}. This individual may either

(i) die with rate μ_i,

(ii) migrate to state $S_{j,n-1}$ $(j\neq i)$ with rate ν_i where it undergoes a similar birth-death-migration process with rates λ_j, μ_j and ν_j respectively,

or (iii) give birth in such a way that as soon as the event of a birth takes place both the individual and its progeny instantly migrate to state $S_{i,n-1}$ where they then both undergo a similar birth-death-migration process with rates λ_i, μ_i and ν_i.

Although Puri follows a slightly different approach, he considers essentially the same subdivision into the set of sub-colonies S_{in}. However, instead of both the individual and its progeny instantly migrating from state S_{in} to $S_{i,n-1}$ in the event of a birth, this time the individual instantly migrates to $S_{i,n-1}$ and its progeny remains in state S_{in}. Equations (3.1) now become modified to

$$(3.4)\quad h_{it}^{(n+1)} = z_i \exp\{-\sigma_i t + \lambda_i \int_o^t h_{is}^{(n)} ds\} + \int_o^t (\mu_i + \nu_i h_{js}^{(n)}) \exp\{-\sigma_i (t-s) + \lambda_i \int_s^t h_{i\tau}^{(n)} d\tau\} ds$$

where $\{h_{it}^{(n)}(\underline{z})\}$ denotes the required sequence of p.g.f's corresponding to Puri's approach. He proves that the sequence of p.g.f's $\{h_{it}^{(n)}\}$ defined by equation (3.4) together with the boundary condition

$$(3.5)\quad h_{it}^{(0)} = 1$$

is monotonic decreasing in n and converges to the solution of the backward equations (2.1).

For any fixed value of n the total number of different realizations possible for both of the first two approaches is strictly bounded above. However, in Puri's situation this is not so. For if an individual gives birth, its progeny remains in the same sub-colony and this could (with probability zero) be repeated indefinitely. We might, therefore, intuitively expect that $h_{it}^{(n)}$ approaches $g_i(t)$ as n tends to infinity at a faster rate than either $f_{it}^{(n)}$ or $g_{it}^{(n)}$.

The functions $f_{it}^{(n)}(\underline{z})$ and $g_{it}^{(n)}(\underline{z})$ monotonically decrease and increase respectively as n successively takes the values $n=0,1,2,\ldots$. Thus if $z_1=z_2=0$ we might expect the function

$$(3.6)\quad \tfrac{1}{2}[f_{it}^{(n)}(\underline{0}) + g_{it}^{(n)}(\underline{0})]$$

to converge reasonably quickly to the extinction probability $g_i(\underline{0};t)$. For example, if $n=1$ expression (3.6) gives

$$(3.7)\quad g_i(\underline{0};t) \simeq \tfrac{1}{2}[1+(\mu_i/\sigma_i)](1-e^{-\sigma_i t}) .$$

We shall show in the next chapter that expression (3.7) is a very poor approximation and, although better approximations may be obtained by taking $n=2,3,4,\ldots$, a sufficiently accurate result would doubtless be too cumbersome to be worthwhile deriving. A similar criticism may be levelled at the p.g.f's $f_{it}^{(n)}(\underline{z})$, $g_{it}^{(n)}(\underline{z})$ and $h_{it}^{(n)}(\underline{z})$. The recursive approach considered in this section should not be thought of as a means of deriving approximations but rather as a way to develop an exact mathematical solution to the backward equations in terms of the limit of convergent sequences of functions.

CHAPTER 6

EXTINCTION PROBABILITIES : TWO-COLONY MODELS

6.1 Introduction

If either immigration rate α_i $(i=1,2)$ is non-zero, both colonies may be empty with non-zero probability at any particular time but extinction is impossible. If $\alpha_1=\alpha_2=0$ ultimate extinction becomes a possibility, and questions to be answered include "When, if ever, is ultimate extinction a certainty?" and "Can an expression be obtained for the probability of extinction by time t?".

Suppress immigration, so that $\alpha_1=\alpha_2=0$, and let $p_i(t)$ $(i=1,2)$ denote the probability of extinction by time t subject to the initial conditions

$$X_i(0) = 1 \quad , \quad X_j(0) = 0 \qquad (i=1,2;\ j\neq i) \quad . \tag{1.1}$$

Then placing $z_1=z_2=0$ in the Kolmogorov backward differential equations (5.2.1) we obtain the equations

$$\begin{aligned} dp_1(t)/dt &= \lambda_1 p_1^2(t) - (\lambda_1+\mu_1+\nu_1)p_1(t) + \nu_1 p_2(t) + \mu_1 \\ dp_2(t)/dt &= \lambda_2 p_2^2(t) - (\lambda_2+\mu_2+\nu_2)p_2(t) + \nu_2 p_1(t) + \mu_2 \quad . \end{aligned} \tag{1.2}$$

Alternatively, we may derive equations (1.2) directly from first principles. For with the initial conditions (1.1), the first event happens in the small time interval $(s,s+ds)$ with probability

$$\sigma_i e^{-\sigma_i s}\, ds$$

where $\sigma_i=\lambda_i+\mu_i+\nu_i$ $(i=1,2)$. Moreover, this event is either a birth, death or migration with probabilities λ_i/σ_i, μ_i/σ_i and

ν_i/σ_i , respectively. Thus on letting s range over all values in $(0,t)$, we have for $i,j=1,2;\ j\neq i$

$$p_i(t) = \int_0^t e^{-\sigma_i s}[\lambda_i p_i^2(t-s) + \nu_i p_j(t-s) + \mu_i]ds \quad .$$

Placing $\tau=t-s$ we may write this as

$$p_i(t)e^{\sigma_i t} = \int_0^t e^{\sigma_i \tau}[\lambda_i p_i^2(\tau)+\nu_i p_j(\tau)+\mu_i]d\tau$$

and this expression differentiates to give the equations

$$e^{\sigma_i t}[\sigma_i p_i(t)+dp_i(t)/dt] = e^{\sigma_i t}[\lambda_i p_i^2(t)+\nu_i p_j(t)+\mu_i]$$

which, on division of both sides by $\exp(\sigma_i t)$, agree with equations (1.2) exactly.

6.2 Probability of Ultimate Extinction

The two simultaneous Riccati differential equations (1.2) do not permit a direct solution, to the best of my knowledge. Even if $\nu_1=0$ (say), so that $p_1(t)$ may be determined, the resulting equation for $p_2(t)$ still has a very complicated form. However, the probability of ultimate extinction, denoted by

$$q_i = \lim_{t\to\infty} p_i(t) \qquad (i=1,2) \ , \tag{2.1}$$

may be found exactly. For on letting $t\to\infty$ in equations (1.2), we have

$$\begin{aligned} 0 &= \lambda_1 q_1^2 - \sigma_1 q_1 + \nu_1 q_2 + \mu_1 \\ 0 &= \lambda_2 q_2^2 - \sigma_2 q_2 + \nu_2 q_1 + \mu_2 \quad , \end{aligned} \tag{2.2}$$

and the required probabilities are given by the smallest solution (q_1,q_2) which lies in the unit square $0 \leq q_1,q_2 \leq 1$.

We may solve equations (2.2) by first noting that $q_1=q_2=1$ is one solution, and then eliminating q_2 (say) so that

$$(q_1-1)(\text{cubic in } q_1) = 0 .$$

This cubic expression may be factored, and the required root subsequently chosen.

Ultimate extinction is certain if and only if $q_1=q_2=1$ is the smallest admissible solution. Puri (1968) considers the geometry of the two parabolas (2.2) and shows that for $q_1=q_2=1$ to be the only admissible solution of equations (2.2) it is necessary and sufficient that

(2.3) $$\xi_i < 0 \quad (i=1,2) \quad \text{and} \quad \xi_1\xi_2 \geq \nu_1\nu_2$$

where $\xi_i=\lambda_i-\mu_i-\nu_i$. In two very recent papers Alksland (1975) and Helland (1975) allow migration between N colonies, instead of just two, and they develop the analogous condition to (2.3).

Puri's condition (2.3) coincides exactly with my previous result when I discussed 'critical migration' in section 2.3.iv. For I showed (result (2.3.38)) that if condition (2.3) holds, the total mean population size $m(t)$ either decreases asymptotically to zero or remains constant. If condition (2.3) does not hold, $m(t)$ increases without bound. If $m(t)$ remains bounded for all $t \geq 0$ we expect stochastic variation to make extinction certain $(q_1=q_2=1)$, whilst if $m(t)\to\infty$ as $t\to\infty$ we expect the actual population size to grow indefinitely large with positive probability, in which case extinction is uncertain $(0 \leq q_1,q_2 < 1)$.

6.3 Approximate Solutions to the Probabilities of Extinction

As equations (1.2) do not possess a simple explicit solution I shall now derive and compare several approximate solutions for the extinction probabilities $p_i(t)$. It follows from the results proved in Chapter 5 that

$$g_{it}^{(n)}(\underline{0}) \leq p_i(t) \leq f_{it}^{(n)}(\underline{0}) \qquad (i=1,2;\ n=0,1,2,\ldots) .$$

On taking $n=1$ we have

$$(\mu_i/\sigma_i)(1-e^{-\sigma_i t}) \leq p_i(t) \leq (1-e^{-\sigma_i t}) \quad , \tag{3.1}$$

so let us first consider the approximation

$$p_i(t) \doteq p_i^{(1)}(t) = q_i(1-e^{-\sigma_i t}) \quad . \tag{3.2}$$

Clearly $p_i^{(1)}(0)=0$ and $p_i^{(1)}(\infty)=q_i$ as required.

If birth and death rates are spatially homogeneous, so that $\lambda_1=\lambda_2=\lambda$ and $\mu_1=\mu_2=\mu$, spatial separation may be ignored whence

$$p_1(t) = p_2(t) = (\mu-\mu e^{-(\lambda-\mu)t})/(\lambda-\mu e^{-(\lambda-\mu)t}) \tag{3.3}$$

(Cox and Miller (1965), p.166). In particular

$$q_1 = q_2 = \begin{cases} 1 & : \ \lambda \leq \mu \\ \mu/\lambda & : \ \lambda > \mu \end{cases} , \tag{3.4}$$

so that for $\lambda > \mu$ expression (3.2) becomes

$$p_i^{(1)}(t) = (\mu/\lambda)(1-e^{-(\lambda+\mu+\nu_i)t}) \qquad (i=1,2) \quad . \tag{3.5}$$

That migration does not affect the extinction probabilities $p_i(t)$ in these circumstances is easily proved analytically. The difference between equations (1.2) is

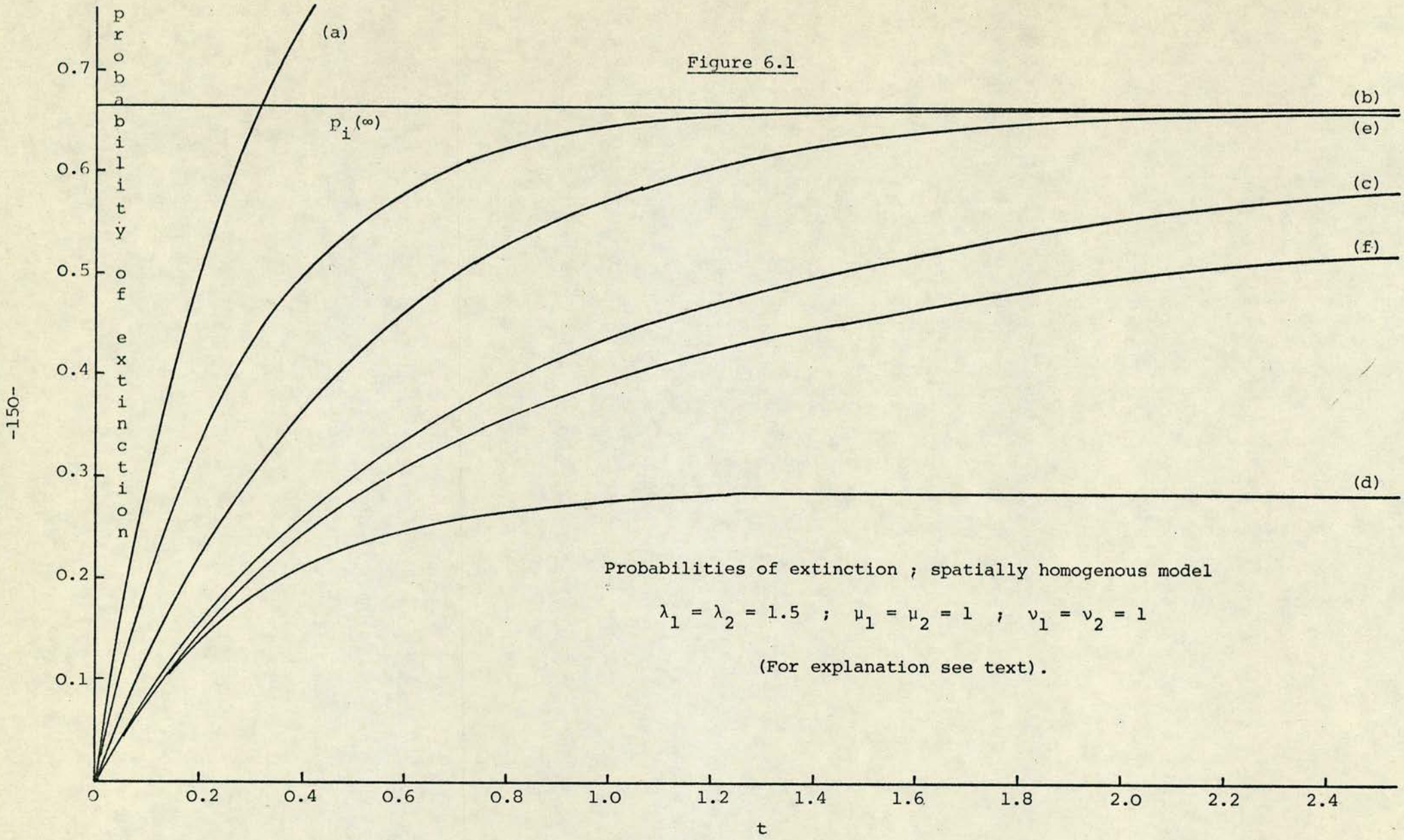


Figure 6.1

Probabilities of extinction ; spatially homogenous model

$\lambda_1 = \lambda_2 = 1.5$; $\mu_1 = \mu_2 = 1$; $\nu_1 = \nu_2 = 1$

(For explanation see text).

$$\frac{d}{dt}[p_1(t)-p_2(t)] = [p_1(t)-p_2(t)]\{\lambda[p_1(t)+p_2(t)] - (\lambda+\mu+\nu_1+\nu_2)\}$$

which integrates to give

$$p_1(t)-p_2(t) = \text{const.}\exp\{\int^t[\lambda(p_1(s)+p_2(s)) - (\lambda+\mu+\nu_1+\nu_2)]ds\} \quad .$$

But at $t=0$

$$p_1(0) = p_2(0) = 0 \quad .$$

Hence const. = 0 and so

(3.6) $$p_1(t) \equiv p_2(t)$$

for all $t \geq 0$. Substituting (3.6) in (1.2) we obtain the differential equation

$$dp_1(t)/dt = \lambda p_1^2(t) - (\lambda+\mu)p_1(t) + \mu$$

and this integrates to give (3.3).

In order to gauge the effectiveness of the approximation (3.5), expressions (3.1), (3.3) and (3.5) are plotted in figure 6.1 for the parameter values $\lambda=1.5$, $\mu=1$ and $\nu_1=\nu_2=1$. Curves (a) and (d) correspond to the upper and lower bounds of the inequality (3.1), curve (b) to the approximation (3.5) and curve (c) to the true solution (3.3). The weighted sum (b) clearly has far too steep a gradient when t is small. In fact

$$\max_{t\geq 0}[p_1^{(1)}(t) - p_1(t)] \doteq 0.24$$

which is unacceptably large. However, in view of the nature of the construction of the sequences $\{f_{it}^{(n)}(\underline{z})\}$ and $\{g_{it}^{(n)}(\underline{z})\}$ in section 5.3, it would seem very unlikely that the first iteration, which generates $f_{it}^{(1)}(\underline{z})$ and $g_{it}^{(1)}(\underline{z})$, could yield good bounds for $g_i(t)$. Hence, it is not surprising that

$$p_i^{(1)}(t) \equiv q_i f_{it}^{(1)}(0,0)$$

is a particularly poor approximation to $p_i(t)$.

An improvement on this approximation might be expected if we work with Puri's sequence $\{h_{it}^{(n)}(\underline{z})\}$, which is generated by expressions (5.3.4) and (5.3.5). For we have already seen at the end of section 5.3 that intuitively we expect the sequence $\{h_{it}^{(n)}\}$ to converge to $g_i(t)$ at a faster rate than either $\{f_{it}^{(n)}\}$ or $\{g_{it}^{(n)}\}$. Combining (5.3.4) and (5.3.5) we have

$$h_{it}^{(1)}(\underline{z}) = 1 - (1-z_i)e^{-(\mu_i+\nu_i)t} \quad . \tag{3.7}$$

As $\{h_{it}^{(n)}\}$ is monotonic decreasing in n , if we put $z_1=z_2=0$ in (3.7) we get

$$0 \leq p_i(t) \leq 1-e^{-(\mu_i+\nu_i)t}$$

which suggests the approximation

$$p_i(t) \doteq p_i^{(2)}(t) = q_i(1-e^{-(\mu_i+\nu_i)t}) \qquad (i=1,2) \quad . \tag{3.8}$$

Expression (3.8) corresponds to curve (e) in figure 6.1 and represents an improvement over the previous approximation (b). In fact

$$\max_{t\geq 0}[p_1^{(2)}(t) - p_1(t)] \doteq 0.14 \quad .$$

Both $p_i^{(1)}(t)$ and $p_i^{(2)}(t)$ are clearly unsatisfactory approximations, and yet if $f_{it}^{(n)}$, $g_{it}^{(n)}$ and $h_{it}^{(n)}$ are generated for values of $n>1$ the resulting approximations for $p_i(t)$ are too cumbersome.

A preferable approach is to look elsewhere for simple

monotonically increasing functions $x_i(t)$ which satisfy $x_i(0)=0$ and $x_i(\infty)=q_i$, and to ensure that at least the first derivatives $x_i'(t)$ and $p_i'(t)$ match at $t=0$. For example,

$$p_i(t) \doteq x_i(t) = \frac{q_i t}{t+(q_i/\mu_i)} \qquad (i=1,2) \tag{3.9}$$

is a ratio of two linear expressions in t which satisfy $x_i(0)=0$, $x_i(\infty)=q_i$ and $x_i'(0)=p_i'(0)$. Expression (3.9) corresponds to curve (f) in figure 6.1, and we see that even this simple linear quotient form is an improvement over (b) and (e). However, only a little extra complication results in an approximation which becomes exact when $\lambda_1=\lambda_2$ and $\mu_1=\mu_2$. For consider the function

$$y_i(t) = q_i(1-e^{b_i t})/(1-c_i e^{b_i t}) \qquad (i=1,2) \tag{3.10}$$

where, without loss of generality, we assume that $b_i < 0$. Clearly $y_i(0)=0$ and $y_i(\infty)=q_i$ as required. We shall choose the constants b_i and c_i so that the first two derivatives of $p_i(t)$ and $y_i(t)$ match at $t=0$.

Differentiating (3.10) and placing $t=0$ we get

$$\begin{aligned} y_i'(0) &= -q_i b_i/(1-c_i) \\ y_i''(0) &= -q_i b_i^2(1+c_i)/(1-c_i)^2 \end{aligned} \tag{3.11}$$

whilst from (1.2) we may obtain

$$\begin{aligned} p_i'(0) &= \mu_i \\ p_i''(0) &= -\mu_i(\lambda_i+\mu_i+\nu_i) + \nu_i\mu_j \end{aligned} \tag{3.12}$$

Placing $y_i'(0)=p_i'(0)$ and $y_i''(0)=p_i''(0)$, and then solving the subsequent equations for b_i and c_i we have

$$b_i = (\lambda_i+\mu_i+\nu_i) - \nu_i(\mu_j/\mu_i) - 2(\mu_i/q_i)$$

(3.13)

$$c_i = -1 + (q_i/\mu_i^2)[\mu_i(\lambda_i+\mu_i+\nu_i)-\nu_i\mu_j] \quad .$$

If $\lambda_1=\lambda_2=\lambda$ and $\mu_1=\mu_2=\mu$ expression (3.10) reduces to (3.3). In these circumstances the approximation $y_i(t)$ and the true value $p_i(t)$ are identical. If the parameter values λ_i, μ_i, ν_i (i=1,2) are known but $\lambda_1 \neq \lambda_2$ and $\mu_1 \neq \mu_2$, equations (1.2) are probably best analysed by using numerical methods.

CHAPTER 7

A SIMULATION STUDY OF THE TWO-COLONY MODEL

7.1 Introduction

Although explicit analytic solutions for the probabilities $p_{ij}(t)$ cannot be derived in general, simulation techniques may be used either to observe realizations of the process or, if they exist, to compute the equilibrium probabilities $p_{ij}(\infty)$. Suppose that the j^{th} event (birth, death, migration or immigration) occurs at time t_j, and let the sizes of the two colonies immediately after the occurrence of this event be $X_1(t_j)$ and $X_2(t_j)$. Then the time $s_j=t_{j+1}-t_j$ to the next event has an exponential distribution with parameter

$$q(t_j) = \sum_{i=1}^{2} [\alpha_i + (\lambda_i+\mu_i+\nu_i)X_i(t_j)] \quad , \tag{1.1}$$

and we shall show that values of the random variable s_j are easily obtained by the use of a suitable pseudo-random number generator. Moreover,

$$\begin{aligned}
&Pr[X_1(t_{j+1})=X_1(t_j)+1 \ , \ X_2(t_{j+1})=X_2(t_j)] = [\alpha_1+\lambda_1 X_1(t_j)]/q(t_j) \\
&Pr[X_1(t_{j+1})=X_1(t_j) \ , \ X_2(t_{j+1})=X_2(t_j)+1] = [\alpha_2+\lambda_2 X_2(t_j)]/q(t_j) \\
&Pr[X_1(t_{j+1})=X_1(t_j)-1 \ , \ X_2(t_{j+1})=X_2(t_j)] = \mu_1 X_1(t_j)/q(t_j) \\
&Pr[X_1(t_{j+1})=X_1(t_j) \ , \ X_2(t_{j+1})=X_2(t_j)-1] = \mu_2 X_2(t_j)/q(t_j) \\
&Pr[X_1(t_{j+1})=X_1(t_j)-1 \ , \ X_2(t_{j+1})=X_2(t_j)+1] = \nu_1 X_1(t_j)/q(t_j) \\
&Pr[X_1(t_{j+1})=X_1(t_j)+1 \ , \ X_2(t_{j+1})=X_2(t_j)-1] = \nu_2 X_2(t_j)/q(t_j) \quad ,
\end{aligned} \tag{1.2}$$

and events corresponding to these probabilities are again easily generated on a computer. Thus both the type of successive events

and their time of occurrence may be computed, which enables us to simulate the development of the two colonies.

The simulation study was carried out on the IBM 370/158 computer at the Edinburgh Regional Computing Centre; the language used was Fortran IV. The pseudo-random numbers were generated successively by the following power-residue method, where YFL denotes the required pseudo-random number in the range [0,1], and RAND is an integer.

```
                RAND = RAND*65539
(1.3)           YFL  = RAND
                YFL  = YFL*.2328307E-9 + .5
```

The generator was started by putting RAND equal to an odd number between 50,000 and 500,000, and the first 100 numbers generated were discarded in order to let the routine settle down.

7.2 Program Description

A. Time to next event : Let $R_0, R_1, \ldots$ be a sequence of pseudo-random numbers generated by the routine (1.3). Then as the time s_j between the j^{th} and $j+1^{th}$ event has an exponential distribution with parameter $q(t_j)$, we may generate the s_j from

$$s_j = -(\log R_j)/q(t_j) \qquad (j=0,1,2,\ldots) \tag{2.1}$$

where $t_0=0$. Thus, for a particular sequence $\{R_i\}$, events occur at times

$$s_0 \ , \ s_0+s_1 \ , \ s_0+s_1+s_2 \ , \ \cdots \ .$$

B. Type of event : Let $L_0, L_1, \ldots$ be a sequence of pseudo-random numbers, uniformly distributed on [0,1], generated by the routine (1.3). Then as the different types of events have the associated

probabilities (1.2), the successive changes in population sizes may be simulated as follows.

Denote $T_j=L_jq(t_j)$ and write $n_1=X_1(t_j)$, $n_2=X_2(t_j)$. Put

$z_1 = \lambda_1 n_1+\alpha_1$	: if $T_j \leq z_1$	$n_1 \to n_1+1$,	$n_2 \to n_2$
$z_2 = z_1+\lambda_2 n_2+\alpha_2$	: if $z_1 < T_j \leq z_2$	$n_1 \to n_1$,	$n_2 \to n_2+1$
$z_3 = z_2+\mu_1 n_1$	: if $z_2 < T_j \leq z_3$	$n_1 \to n_1-1$,	$n_2 \to n_2$
$z_4 = z_3+\mu_2 n_2$	: if $z_3 < T_j \leq z_4$	$n_1 \to n_1$,	$n_2 \to n_2-1$
$z_5 = z_4+\nu_1 n_1$	: if $z_4 < T_j \leq z_5$	$n_1 \to n_1-1$,	$n_2 \to n_2+1$
	otherwise	$n_1 \to n_1+1$,	$n_2 \to n_2-1$.

On commencing with $X_i(0)=a_i$ $(i=1,2)$ and letting j take the successive values $j=0,1,2,\ldots$, this procedure simulates the successive states of the process.

C. Output : Routines A and B may now be used to obtain realizations of the process, examples of which are portrayed in figures 2.2a, 2.2b and 2.2c. In these diagrams the state of the process is plotted every 0.1 time units (except for the opening stages in 2.2c). However, as is mentioned in section 2.2, many events may occur in each time period of length 0.1 and for these examples the actual behaviour of the process is far more 'spiky' than is suggested by the figures. As an example, the realization corresponding to figure 2.2a commences as follows.

Event	Time (t_j)	$X_1(t_j)$	$X_2(t_j)$	Type of event
0	0	6	14	
1	0.0117	7	13	migration $2 \to 1$
2	0.0310	7	14	increase in 2
3	0.0345	7	15	increase in 2
4	0.0435	6	15	death in 1
5	0.0448	6	16	increase in 2
6	0.0560	7	16	increase in 1
7	0.0617	7	15	death in 2
⋮	⋮	⋮	⋮	⋮

An artificial realization of a simple one-colony birth-death process is contained in Kendall (1950).

D. Equilibrium probabilities : If $\omega_1 < 0$ an equilibrium distribution exists and the above procedure is easily extended to generate values for the equilibrium probabilities $p_{ij}(\infty)$. Let a particular realization (say the n^{th}) last for $T^{(n)}$ time units and let $T_{ij}^{(n)}$ be the total length of time for this realization in which colony 1 is of size i and colony 2 is of size j . Then for this particular realization the equilibrium probabilities may be 'estimated' by

$$q_{ij}^{(n)}(\infty) = T_{ij}^{(n)}/T^{(n)} \qquad (i,j=0,1,\ldots,L) \; . \tag{2.2}$$

The size of the probability matrix to be stored in the computer has to be declared at the start of the program. Hence a finite upper bound L is chosen, yielding a $(L+1)\times(L+1)$ matrix, which is large enough to ensure that states corresponding to either i or j greater than L are extremely unlikely to be reached in

the duration of a realization. All the simulations described in this chapter were performed with L=49. Any non-zero values of $T_{ij}^{(n)}$, which had i or $j \geq 50$, that did occur were aggregated in the form

$$X^{(n)} = \sum_{i \text{ or } j \geq 50} T_{ij}^{(n)} ,$$

and the value of $X^{(n)}/T^{(n)}$ was printed on the output to check that L had been chosen sufficiently large.

Suppose that we replicate the simulation N times (n=1,...,N) by choosing different starting values for the initial random number. The probabilities (2.2) may then be combined to yield the 'estimator'

$$\hat{p}_{ij}(\infty) = (1/N) \sum_{n=1}^{N} q_{ij}^{(n)}(\infty) , \tag{2.3}$$

and the variance of the N values of $q_{ij}^{(n)}(\infty)$ (n=1,...,N) yields a measure of the accuracy of the result.

In order to illustrate this technique let us consider the following example. A total of 12 different realizations were computed for the process with the following parameter values.

	λ_i	μ_i	ν_i	α_i	a_i
Colony 1	2.0	3.5	0.3	1.0	5
Colony 2	1.5	2.3	5.0	2.4	1

Each migration consists of 60,000 events. The probabilities $q_{ij}^{(n)}(\infty)$ (n=1,...,12), together with their first- and second-order moments, were obtained for each realization. The simulated probabilities $\hat{p}_{ij}(\infty)$ were then derived from (2.3) with N=12 .

Their values to 4 decimal places are shown in table 7.1 for $i=0,\ldots,7$ and $j=0,\ldots,6$.

Table 7.1

Values of the simulated probabilities $\hat{p}_{ij}(\infty)$, $\hat{p}_{i.}(\infty)$ and $\hat{p}_{.j}(\infty)$

(For model see text)

i \ j	0	1	2	3	4	5	6	$\hat{p}_{i.}(\infty)$
0	0.1976	0.0667	0.0174	0.0041	0.0009	0.0002	0.0001	0.2870
1	0.1529	0.0593	0.0173	0.0046	0.0012	0.0003	0.0001	0.2357
2	0.1054	0.0443	0.0138	0.0040	0.0010	0.0003	0.0001	0.1689
3	0.0683	0.0311	0.0103	0.0031	0.0008	0.0002	0.0001	0.1140
4	0.0429	0.0206	0.0072	0.0022	0.0007	0.0002	0.0001	0.0738
5	0.0265	0.0134	0.0048	0.0016	0.0005	0.0001	0.0000	0.0470
6	0.0160	0.0085	0.0033	0.0010	0.0003	0.0001	0.0000	0.0292
7	0.0095	0.0053	0.0020	0.0007	0.0002	0.0001	0.0000	0.0178
$\hat{p}_{.j}(\infty)$	0.6325	0.2574	0.0796	0.0225	0.0060	0.0016	0.0004	

The simulated marginal probabilities $\hat{p}_{i.}(\infty)$ and $\hat{p}_{.j}(\infty)$ denote the values

$$\hat{p}_{i.}(\infty) = \sum_{j=0}^{L} \hat{p}_{ij}(\infty) \quad \text{and} \quad \hat{p}_{.j}(\infty) = \sum_{i=0}^{L} \hat{p}_{ij}(\infty) \quad .$$

This program was also used to check the validity of expressions (2.3.24) and (2.4.21)-(2.4.23) for the first- and second-order moments $m_i(\infty)$ and $V_{ij}(\infty)$ $(i,j=1,2)$. The process was simulated for several different sets of parameter values; for each set the 12 replications were used to compute the simulated mean and standard deviation for each of the 5 moments. In every case the simulated

values were within one standard deviation of the theoretical values calculated from (2.3.24) and (2.4.21)-(2.4.23).

As an illustration, let us consider the example described above. The first two columns in table 7.2 give the simulated and theoretical moments, the third column their difference and the fourth column gives the standard deviation computed from the 12 replicates. The number of time units elapsed during each of the 12 realizations of 60,000 events has a mean value of 3071.29 and a standard deviation of 54.67.

Table 7.2

A comparison between the simulated and the theoretical first- and second-order moments

(For model see text)

	Simulated	Theoretical	Difference	S. Deviation
$m_1(\infty)$	1.9995	1.9911	0.0084	0.0548
$m_2(\infty)$	0.5194	0.5168	0.0026	0.0097
$V_{11}(\infty)$	4.7299	4.7718	-0.0419	0.1941
$V_{12}(\infty)$	0.2018	0.2047	-0.0029	0.0151
$V_{22}(\infty)$	0.6664	0.6610	0.0054	0.0174

7.3 A Comparison of Two Approximations to Equilibrium Probabilities

In Chapter 4 I developed the modified process and the binomial approximation, and showed that in general there are fairly strong reasons for favouring the latter as the better approximation to the

true equilibrium distribution of population size. Moreover, if $\nu_1=\nu_2=0$ the negative binomial probabilities are exact, and so it is reasonable to expect them to be relatively close to the true probabilities for sufficiently small values of ν_1 and ν_2.

To illustrate the effect of the magnitude of ν_1 and ν_2 on the goodness-of-fit of these two approximate distributions, the process was simulated (as described in the previous section) for various values of ν_1 and ν_2. For simplicity we restrict our attention to the spatially homogeneous model and consider only the equilibrium situation. The birth, death and immigration parameters are given the fixed values

$$\lambda = 1.0 \quad ; \quad \mu = 1.2 \quad ; \quad \alpha = 0.3$$

respectively, and the migration rates take the values

$$\nu = 0 \ , \ 0.1 \ \text{and} \ 5.0$$

on successive simulations.

In equilibrium the modified process behaves as two independent Poisson processes, the joint probabilities being given by

$$(4.2.15)\quad p_{ij}(\infty) = [m_1(\infty)]^i [m_2(\infty)]^j (i!j!)^{-1} \exp\{-[m_1(\infty)+m_2(\infty)]\} \ ,$$

whilst the probabilities associated with the negative binomial approximation are given by (4.3.2). Both of these expressions are easy to compute numerically. It follows from expression (2.3.24) that

$$m_1(\infty) = m_2(\infty) = \alpha/(\mu-\lambda) \qquad (\mu > \lambda) \ .$$

This expression does not contain the parameter ν, and so the set of probabilities (4.2.15) is completely unaffected by changes in the migration rate.

Table 7.3 shows the probabilities $p_{ij}(\infty)$ $(i,j=0,\ldots,7)$ corresponding to $\nu=0$. In each vertical grouping of three the top, middle and bottom numbers refer to the simulated, negative binomial and Poisson probabilities respectively. In this situation the negative binomial probabilities are exact (see result (4.3.14)) and so a comparison between the simulated and negative binomial probabilities provides some measure of the accuracy of the simulation results. When $i=j=0$ the absolute error is 0.0039 (this is 1.14% too low), and over the remaining values $(i,j \geq 0,\ i+j \neq 0)$ the largest absolute error is 0.0011 (this is 1.29% too low) which occurs when $i=1,\ j=0$. To compensate for these two low values most of the other probabilities are slightly inflated.

Tables 7.4 and 7.5 show the probabilities $p_{ij}(\infty)$ $(i,j=0,\ldots,7)$ corresponding to $\nu=0.1$ and $\nu=5$, respectively. In each vertical grouping of two the top number is the average simulated probability, given by $\frac{1}{2}(p_{ij}(\infty) + p_{ji}(\infty))$, and the bottom number is the corresponding negative binomial approximation. Only the upper triangles are shown as the probability matrices are symmetrical. The Poisson probabilities are the same as those given in table 7.3.

A visual inspection of tables 7.3-7.5 shows that the Poisson approximation is hopelessly inadequate in this situation. However, it should not be forgotten that, unlike its negative binomial counterpart, it does at least have a physical interpretation. The failure of the Poisson probabilities as satisfactory approximations is mainly due to the overshadowing effect of the deterministic births, which are equivalent to an extra immigration component. This causes a considerable reduction in the spread of the

Table 7.3

A comparison of the simulated, negative binomial and Poisson approximations to the equilibrium probabilities $p_{ij}(\infty)$: $\nu=0$

$\lambda = 1.0$; $\mu = 1.2$; $\alpha = 0.3$

0	1	2	3	4	5	6	7	i
0.3374	0.0851	0.0464	0.0291	0.0203	0.0149	0.0111	0.0082	j
0.3413	0.0853	0.0462	0.0295	0.0203	0.0146	0.0107	0.0080	0
0.0498	0.0747	0.0560	0.0280	0.0105	0.0032	0.0008	0.0002	
0.0842	0.0219	0.0121	0.0076	0.0052	0.0037	0.0027	0.0020	
0.0853	0.0213	0.0116	0.0074	0.0051	0.0036	0.0027	0.0020	1
0.0747	0.1120	0.0840	0.0420	0.0158	0.0047	0.0012	0.0003	
0.0463	0.0120	0.0069	0.0046	0.0030	0.0021	0.0015	0.0011	
0.0462	0.0116	0.0063	0.0040	0.0028	0.0020	0.0015	0.0011	2
0.0560	0.0840	0.0630	0.0315	0.0118	0.0035	0.0009	0.0002	
0.0299	0.0077	0.0042	0.0028	0.0019	0.0013	0.0009	0.0007	
0.0295	0.0074	0.0040	0.0026	0.0018	0.0013	0.0009	0.0007	3
0.0280	0.0420	0.0315	0.0158	0.0059	0.0018	0.0004	0.0001	
0.0205	0.0052	0.0028	0.0018	0.0012	0.0009	0.0007	0.0005	
0.0203	0.0051	0.0028	0.0018	0.0012	0.0009	0.0006	0.0005	4
0.0105	0.0158	0.0118	0.0059	0.0022	0.0007	0.0002	0.0000	
0.0148	0.0037	0.0021	0.0014	0.0009	0.0007	0.0005	0.0003	
0.0146	0.0036	0.0020	0.0013	0.0009	0.0006	0.0005	0.0003	5
0.0032	0.0047	0.0035	0.0018	0.0007	0.0002	0.0001	0.0000	
0.0107	0.0027	0.0015	0.0010	0.0007	0.0005	0.0004	0.0003	
0.0107	0.0027	0.0015	0.0009	0.0006	0.0005	0.0003	0.0003	6
0.0008	0.0012	0.0009	0.0004	0.0002	0.0001	0.0000	0.0000	
0.0077	0.0021	0.0011	0.0008	0.0005	0.0004	0.0003	0.0002	
0.0080	0.0020	0.0011	0.0007	0.0005	0.0003	0.0003	0.0002	7
0.0002	0.0003	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	

In each cell : top number - simulated
middle number - negative binomial
bottom number - Poisson

Table 7.4

A comparison of the simulated and negative binomial approximations to the equilibrium probabilities $p_{ij}(\infty)$: $\nu=0.1$

$\lambda = 1.0$; $\mu = 1.2$; $\alpha = 0.3$

0	1	2	3	4	5	6	7	i
0.3408	0.0849	0.0428	0.0249	0.0157	0.0101	0.0065	0.0045	j
0.2875	0.0908	0.0502	0.0317	0.0213	0.0148	0.0105	0.0076	0
	0.0282	0.0163	0.0107	0.0073	0.0052	0.0037	0.0025	
	0.0287	0.0158	0.0100	0.0067	0.0047	0.0033	0.0024	1
		0.0099	0.0067	0.0046	0.0034	0.0024	0.0018	
		0.0088	0.0055	0.0037	0.0026	0.0018	0.0013	2
			0.0044	0.0032	0.0023	0.0017	0.0013	
			0.0035	0.0023	0.0016	0.0012	0.0008	3
				0.0023	0.0017	0.0013	0.0010	
				0.0016	0.0011	0.0008	0.0006	4
					0.0013	0.0010	0.0007	
					0.0008	0.0005	0.0004	5
						0.0007	0.0006	
						0.0004	0.0003	6
							0.0004	
							0.0002	7

In each cell : top number - simulated
bottom number - negative binomial

Table 7.5

A comparison of the simulated and negative binomial approximations to the equilibrium probabilities $p_{ij}(\infty)$: $\nu=5$

$\lambda = 1.0$; $\mu = 1.2$; $\alpha = 0.3$

0	1	2	3	4	5	6	7	i
0.3403	0.0858	0.0304	0.0116	0.0046	0.0018	0.0008	0.0003	j
0.2252	0.0952	0.0543	0.0337	0.0217	0.0143	0.0096	0.0065	0
	0.0538	0.0292	0.0156	0.0078	0.0038	0.0018	0.0009	
	0.0402	0.0230	0.0142	0.0092	0.0060	0.0040	0.0027	1
		0.0220	0.0140	0.0083	0.0047	0.0026	0.0013	
		0.0131	0.0081	0.0052	0.0035	0.0023	0.0016	2
			0.0107	0.0074	0.0047	0.0029	0.0017	
			0.0050	0.0032	0.0021	0.0014	0.0010	3
				0.0056	0.0041	0.0028	0.0019	
				0.0021	0.0014	0.0009	0.0006	4
					0.0033	0.0025	0.0018	
					0.0009	0.0006	0.0004	5
						0.0020	0.0016	
						0.0004	0.0003	6
							0.0013	
							0.0002	7

In each cell : top number - simulated
bottom number - negative binomial

distribution, as can be seen from the tables.

As this particular model is spatially homogeneous ($\lambda_1 = \lambda_2 = 1.0$, $\mu_1 = \mu_2 = 1.2$ and $\alpha_1 = \alpha_2 = 0.3$), the distribution of the total population size is given by a simple birth-death-immigration process. Thus the summations

$$u_n = \sum_{i=0}^{n} p_{i,n-i}(t) \qquad (n=0,1,2,\ldots)$$

are invariant under changes of ν_1 and ν_2. Note that this necessarily implies that $p_{00}(t)$ and $p_{01}(t) = p_{10}(t)$ are also invariant. The negative binomial values at $t=\infty$ of $u_0,\ldots,u_7$ are given in table 7.6 for the three cases $\nu = 0$, 0.1 and 5, and they clearly change quite considerably as ν increases. When $\nu = 0.1$ and 5 the value of u_0 is too low whilst the values of

Table 7.6

Values of $u_n = \sum_{i=0}^{n} p_{i,n-i}(\infty)$ for $\nu = 0$, 0.1 and 5 :

negative binomial approximation

n	$\nu = 0$	$\nu = 0.1$	$\nu = 5$
0	0.3413	0.2875	0.2252
1	0.1706	0.1876	0.1904
2	0.1137	0.1291	0.1488
3	0.0822	0.0950	0.1134
4	0.0617	0.0714	0.0849
5	0.0474	0.0540	0.0632
6	0.0368	0.0413	0.0466
7	0.0290	0.0316	0.0344

$u_1,\ldots,u_7$ are too high. It appears that for this set of

parameters the negative binomial approximation will be accurate only when ν is considerably less than 0.1. The invariance property of $p_{00}(\infty)$ and $p_{01}(\infty)$ could be used as a guide to the choice of the range of ν over which the approximation is sufficiently good.

Having discussed changes between the u_n, we shall now compare the simulated and negative binomial distributions within the u_n by considering the probabilities

$$r_i^{(n)} = p_{i,n-i}(\infty)/u_n \qquad (i=0,\ldots,n;\ n \geq 0) \ .$$

The values of $r_i^{(n)}$ are shown in figure 7.1 for $n=6$, and it is clear that an increase in ν has a completely different effect on the two distributions. The negative binomial graph remains U-shaped but becomes slightly flatter, whilst the simulated graph undergoes a complete inversion.

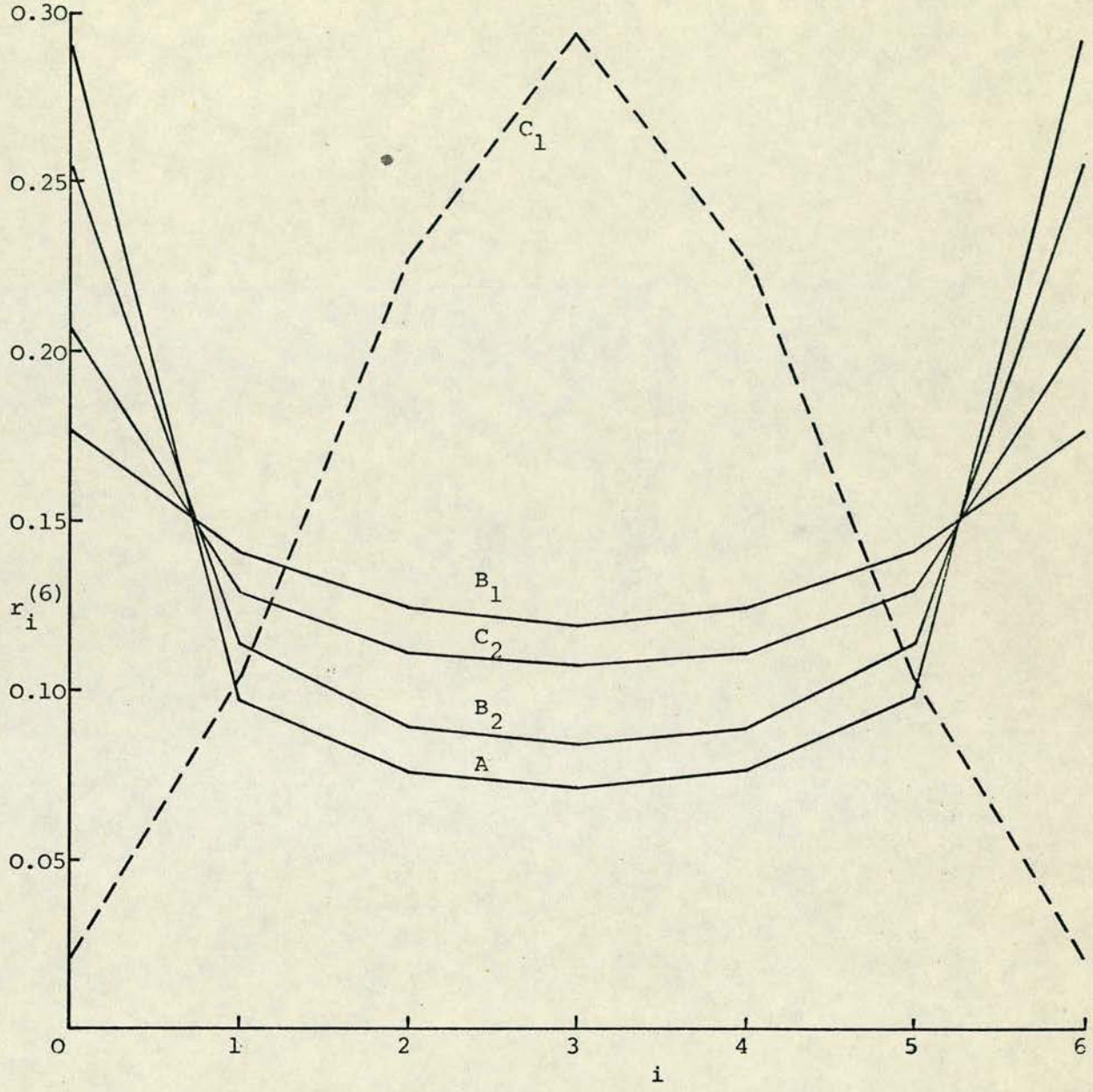
The effect of an increase in migration rate on the simulated marginal probabilities $\hat{p}_{i.}(\infty)$ and $\hat{p}_{.j}(\infty)$ is also quite marked. As ν_1 and ν_2 are allowed to increase the correlation coefficient $V_{12}(\infty)/[V_{11}(\infty)V_{22}(\infty)]^{\frac{1}{2}}$ also increases and this will tend to reduce the spread of the individual population sizes. For if $X_1(\infty)$ (say) wanders too far above its mean value $m_1(\infty)$, the migration parameter ν_1 will act as though it were an added death rate; whilst if $X_1(\infty)$ approaches too close to zero, the other migration parameter ν_2 will act as though it were an added immigration rate. This is illustrated in table 7.7 which shows the first 18 marginal probabilities $\hat{p}_{i.}(\infty)$, corresponding to $\nu=0$ and $\nu=5$, for the simulation example described above ($\lambda_1=\lambda_2=1.0$, $\mu_1=\mu_2=1.2$ and

$\alpha_1 = \alpha_2 = 0.3)$. Here the marginal probabilities for $\nu = 5$ are greater than those for $\nu = 0$ if and only if $1 \leq i \leq 7$.

Table 7.7

A comparison of the simulated marginal probabilities $\hat{p}_{i.}(\infty)$: $\nu = 0$ and $\nu = 5$

i	$\nu = 0$	$\nu = 5$	i	$\nu = 0$	$\nu = 5$	i	$\nu = 0$	$\nu = 5$
0	0.5807	0.4758	6	0.0186	0.0201	12	0.0037	0.0022
1	0.1464	0.1994	7	0.0137	0.0138	13	0.0029	0.0016
2	0.0808	0.1139	8	0.0103	0.0095	14	0.0023	0.0011
3	0.0516	0.0707	9	0.0079	0.0065	15	0.0019	0.0008
4	0.0353	0.0452	10	0.0062	0.0046	16	0.0015	0.0006
5	0.0254	0.0299	11	0.0048	0.0032	17	0.0012	0.0004



A - Simulated/Negative Binomial $\nu = 0$

B_1 - Simulated

B_2 - Negative Binomial $\nu = 0.1$

C_1 - Simulated

C_2 - Negative Binomial $\nu = 5$

Figure 7.1

Values of $r_i^{(6)} = p_{i,6-i}(\infty)/u_n$ for $\nu = 0$, 0.1 and 5

$\lambda_1 = \lambda_2 = 1.0$; $\mu_1 = \mu_2 = 1.2$; $\alpha_1 = \alpha_2 = 0.3$

(For explanation see text).

CHAPTER 8

STEPPING-STONE MODELS FOR MULTIPLE COLONIES

8.1 Introduction

The results so far derived in this thesis have been based on a two-colony population model. In the remaining four chapters we shall consider the effect of migration over more general spatial territories.

In order to investigate the effect of migration between separate geographic regions on population size, Bailey (1968) considered the so-called 'stepping-stone' model of population structure. The model is as follows. The population is composed of an infinite number of colonies situated at the integer points of a single co-ordinate axis represented by $-\infty < i < \infty$. The number of individuals in colony i at time t is denoted by $X_i(t)$ with

$$X_i(0) = a_i \quad . \tag{1.1}$$

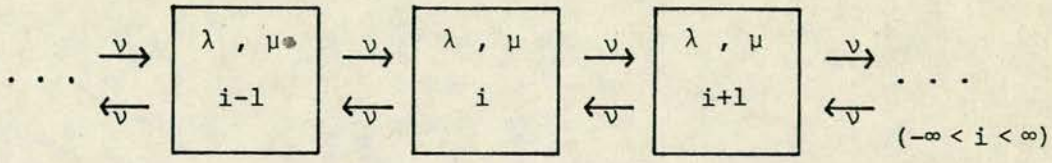
Each colony is considered to be subject to a simple stochastic birth-death-migration process with common birth and death rates λ and μ respectively, and with common migration rates ν to each of the two nearest neighbours $i-1$ and $i+1$. This particular process is illustrated in figure 8.1.

Let the vectors $\underline{x} = (\ldots,x_i,\ldots)$ and $\underline{z} = (\ldots,z_i,\ldots)$ represent two infinite-dimensional vectors with x_i and z_i in the i^{th} positions $(-\infty < i < \infty)$. Define

$$p(\underline{x};t) = \Pr[X_i(t)=x_i,\ -\infty < i < \infty \mid X_i(0)=a_i,\ -\infty < i < \infty]$$

Figure 8.1

Bailey's stepping-stone model for multiple colonies



and let $G(\underline{z};t)$ be the corresponding probability generating function with $0 \leq z_i \leq 1$. Following the 'random-variable technique' described by Bailey ((1964), sections 7.4, 10.1), Bailey (1968) constructs the Kolmogorov forward differential equation for $G(\underline{z};t)$ which is given by

$$(1.2)\quad \frac{\partial G}{\partial t} = \sum_{i=-\infty}^{\infty} [\lambda(z_i^2 - z_i) + \mu(1-z_i) + \nu(z_{i+1} - z_i) + \nu(z_{i-1} - z_i)] \frac{\partial G}{\partial z_i} \quad .$$

Although this equation may not be solved explicitly, he uses it to obtain the first-order moments of the process and the generating function of the second-order moments. He shows that the mean number of individuals in colony i at time t is

$$(1.3)\qquad m_i(t) = e^{(\lambda-\mu-2\nu)t} \sum_{j=-\infty}^{\infty} a_j I_{i-j}(2\nu t) \qquad (-\infty < i < \infty) \quad .$$

Here

$$(1.4)\qquad I_i(x) = (\tfrac{1}{2}x)^i \sum_{k=0}^{\infty} \frac{(\tfrac{1}{4}x^2)^k}{k!(i+k)!}$$

(Abramowitz and Stegun (1965), result (9.6.10)) denotes the modified Bessel function of the first kind. In this chapter we shall first obtain an approximate solution for $G(\underline{z};t)$ and then examine the process when immigration is introduced into a single colony from outside the system. Bailey extends result (1.3) to

cover populations distributed over the nodes of a square or a cubic lattice. In these circumstances an individual is allowed to migrate to one of the four or six nearest neighbours, respectively.

Figure 8.2 shows a simulated realization of such a one-dimensional nearest-neighbour process, which starts at time $t=0$ with one individual sited in colony 0. The parameters take the values $\lambda=\nu=1$ and $\mu=0$, and the figure shows the values of $\log_e(1+X_i(t))$ $(-17 \leq i \leq 17)$ at times $t=0$ (dotted line), 3, 6, 9 and 12. When t is small migration of the one or two individuals present may cause the 'centre' of the distribution to move slightly (to colony -1 in the example given), but as t increases further the 'centre' appears to become fixed with the shape of the distribution developing a 'bullet-nosed' profile. Details of how the simulation is performed are given in section 9.8 for the one-way migration model, and the program is easily augmented to cover the present model in which migration may occur in either direction.

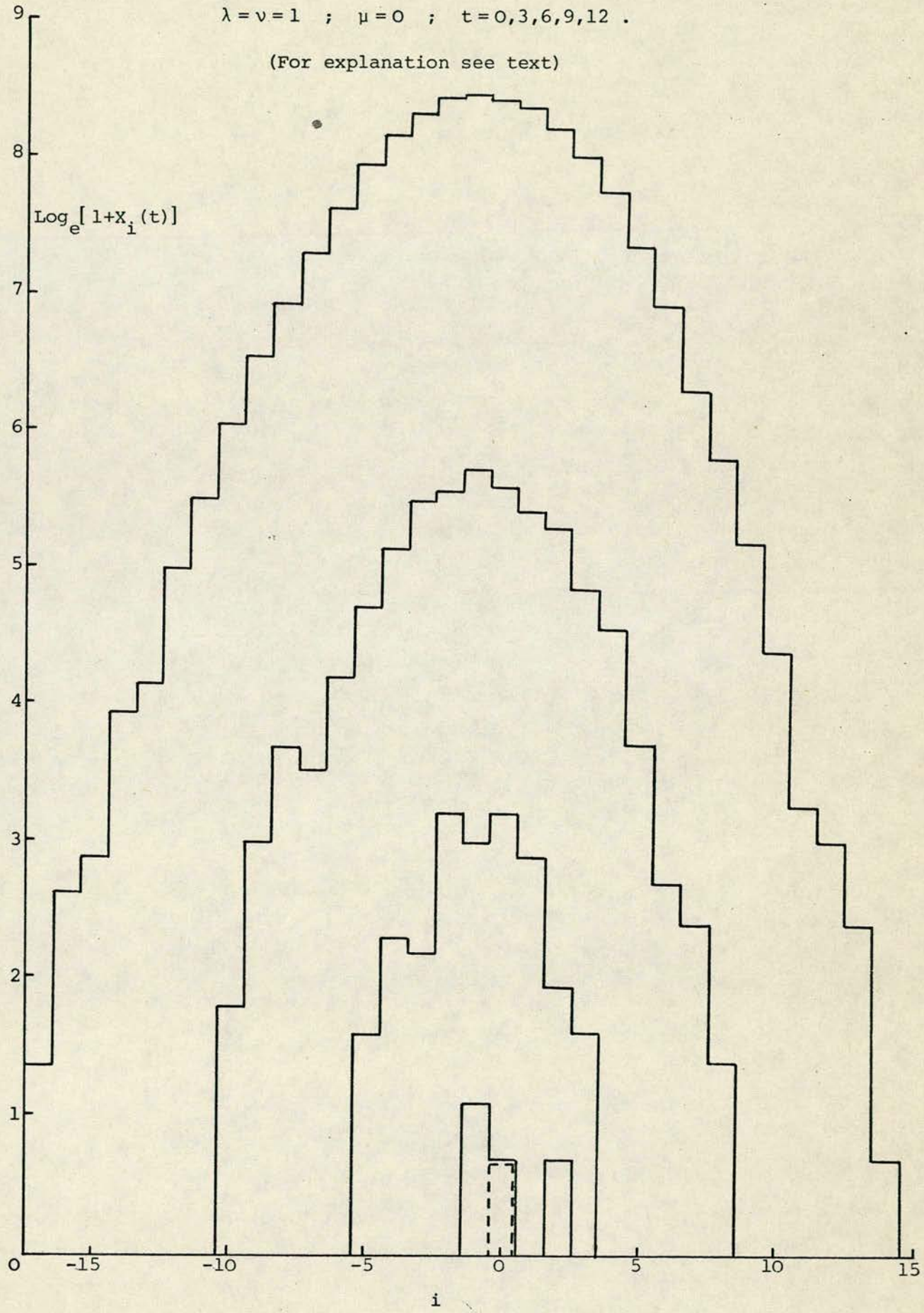
Stepping-stone models have recently been used by several authors in various fields of research. Adke (1969) generalizes Bailey's model by allowing the birth and death rates to be time-dependent. He obtains the first- and second-order moments and derives an iterative solution for the generating function of the colony sizes. Bailey (1968) mentions the desirability of obtaining the joint distribution of the number of individuals in each colony, and Adke (1969) indicates how this may be achieved by

Figure 8.2

A simulation example of Bailey's Migration Process

$\lambda = \nu = 1$; $\mu = 0$; $t = 0,3,6,9,12$.

(For explanation see text)



using the moment generating functional together with Theorem 5.1 in Harris (1963). Adke and Moyal (1963) consider an analogous situation to Bailey's process by replacing migration over the integers with a diffusion process on a line. They develop an iterative procedure for evaluating the generating function of the colony sizes (this parallels Adke (1969)), and discuss asymptotic properties conditional on the population being of size n. This model is generalized by Adke (1964a) to include time-dependent birth and death rates.

If the stepping-stone model is 'truncated', so that there are only a finite number of colonies, 'edge-effects' will clearly occur at the boundaries. Renshaw (1970, 1972) determines the first-order moments in these circumstances and Usher and Williamson (1970) analyse a similar model in discrete time. They consider the population to be split into migrants and nonmigrants, each group having different birth and death rates. Adke (1964b) replaces diffusion over the whole real line by diffusion on a finite interval with reflecting barriers.

Davis (1965, 1967a, 1967b) discusses the general 'branching-diffusion' process in which the population multiplies according to the Bellman-Harris age-dependent branching process and diffuses through a region which has no absorbing barriers. Davis (1970) applies these results to Bailey's model; when there is only a single ancestor Bailey's model is a special case of this 'branching-diffusion' process. Crump (1970) studies a general age-dependent branching process in which the population is distributed in N colonies with migration between them, and he obtains asymptotic

expressions for the first- and second-order moments in several special cases.

The stepping-stone model has also been used in genetics. Kimura (1953) proposed it in order to investigate the effect of the breeding structure of a population on genetic differentiation within the population. It is assumed that the population consists of infinitely many colonies situated at the grid points of an n-dimensional lattice. Migration occurs between nearest neighbours in each generation, and each colony also receives immigrants from the mixture of the whole population in which the gene frequency under consideration is assumed constant in time. Kimura called the former type of migration the short range migration and the latter type the long range migration. The biological implications are discussed in Kimura and Weiss (1964), whilst Weiss and Kimura (1965) obtain results for the genetic correlation and variance between colonies for general cases of the model, assuming that the short range migration is symmetrical in each fixed direction. Maruyama (1969) extends the theory to cases in which the short range migration rates are not fixed in each direction and this is further extended in Maruyama (1970) who analyses stepping-stone models of finite length.

8.2 An Approximate Solution to the Kolmogorov Forward Equation for a Model with no Immigration

8.2.i A general approximation to the probability generating function

Although equation (1.2) has so far evaded a solution in closed

form, we may obtain a good approximation by slightly modifying the birth mechanism. This procedure is equivalent to that used for the two-colony model in section 4.2. We shall take the probability of a birth in colony k in the small time interval $(t,t+\delta t)$ to be $\lambda m_k(t)\delta t+o(\delta t)$ instead of $\lambda x_k \delta t+o(\delta t)$. That is, we replace the random variable x_k by its mean value $m_k(t)$.

Let the infinite-dimensional vector $\underline{x} = (\ldots,x_k,\ldots)$ denote the number of individuals in each colony. Then consideration of all possible events in the small time interval $(t,t+\delta t)$ yields the differential equation

$$\begin{aligned} dp(\underline{x};t)/dt = \sum_{k=-\infty}^{\infty} \{&\lambda m_k(t)p(\underline{x}-\underline{e}_k;t) + \mu(x_k+1)p(\underline{x}+\underline{e}_k;t) \\ &+ \nu(x_{k+1}+1)p(\underline{x}+\underline{e}_{k+1}-\underline{e}_k;t) + \nu(x_{k-1}+1)p(\underline{x}+\underline{e}_{k-1}-\underline{e}_k;t) \\ &- [\lambda m_k(t) + (\mu+2\nu)x_k]p(\underline{x};t)\} \quad , \end{aligned} \tag{2.1}$$

where $\underline{e}_k = (\ldots,0,1,0,\ldots)$ denotes the infinite-dimensional vector with 1 in the k^{th} place and zeros elsewhere. We obtain the Kolmogorov forward differential equation corresponding to (1.2) for the modified p.g.f.

$$G(\underline{z};t) = \sum_{\underline{x}} p(\underline{x};t) \prod_{j=-\infty}^{\infty} z_j^{x_j}$$

by multiplying both sides of equation (2.1) by $\prod_{j=-\infty}^{\infty} z_j^{x_j}$ and summing the resulting expression over $x_j = 0,1,2,\ldots;\ -\infty < j < \infty$. This procedure gives

$$\begin{aligned} \frac{\partial G}{\partial t} = &\sum_{k=-\infty}^{\infty} \lambda m_k(t)(z_k-1)G \\ &+ \sum_{k=-\infty}^{\infty} [\mu(1-z_k) + \nu(z_{k+1}-z_k) + \nu(z_{k-1}-z_k)]\frac{\partial G}{\partial z_k} \end{aligned} \tag{2.2}$$

with

$$G(\underline{z};0) = \prod_{j=-\infty}^{\infty} z_j^{a_j} \quad . \tag{2.3}$$

In order to solve (2.2) we first write down the auxiliary equations

$$\frac{dt}{1} = \frac{-dz_k}{\mu(1-z_k)+\nu(z_{k-1}-z_k)+\nu(z_{k+1}-z_k)} = \frac{dG}{\sum_k \lambda m_k(t)(z_k-1)G} \tag{2.4}$$

$(-\infty < k < \infty)$, and then make the transformation

$$r_k = 1-z_k \qquad (-\infty < k < \infty) \quad ,$$

so that the first set of equations in (2.4) becomes

$$dr_k/dt = \mu r_k + \nu(r_k-r_{k-1}) + \nu(r_k-r_{k+1}) \qquad (-\infty < k < \infty). \tag{2.5}$$

Define the generating function

$$H(\underline{r};y) = \sum_{k=-\infty}^{\infty} r_k y^k \quad , \tag{2.6}$$

where $\underline{r} = (\ldots,r_k,\ldots)$ is an infinite-dimensional vector with r_k in the k^{th} position, and assume the existence of expression (2.6) over some appropriately chosen complex domain of y . Multiplying equations (2.5) by y^k and summing over the range $-\infty < k < \infty$ we get the differential equation

$$\partial H(\underline{r};y)/\partial t = H(\underline{r};y)[(\mu+2\nu)-\nu(y+y^{-1})] \tag{2.7}$$

which integrates to

$$H(\underline{r};y) = \text{const.}\exp\{t[(\mu+2\nu) - \nu(y+y^{-1})]\} \quad . \tag{2.8}$$

As equation (2.7) has been integrated with respect to t , let us write the constant of integration in the general form

$$\sum_{k=-\infty}^{\infty} A_k y^k$$

where $A_k (-\infty < k < \infty)$ are constants. Using the general expansion for modified Bessel functions of the first kind given by

$$\exp\{x(y+y^{-1})\} = \sum_{j=-\infty}^{\infty} I_j(2x)y^j \qquad (y \neq 0) \tag{2.9}$$

(Abramowitz and Stegun (1965), result (9.6.33)), we see that expression (2.8) may now be written in the form

$$H(\underline{r};y) = e^{(\mu+2\nu)t} \sum_{k=-\infty}^{\infty} A_k y^k \sum_{j=-\infty}^{\infty} I_j(-2\nu t)y^j \quad . \tag{2.10}$$

Extracting the coefficient of y^k from (2.10) we obtain

$$r_k = e^{(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} A_{k-i} I_i(-2\nu t)$$

and so

$$z_k = 1-e^{(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} A_{k-i} I_i(-2\nu t) \quad . \tag{2.11}$$

Similarly, if we write expression (2.10) in the form

$$\exp\{\nu t(y+y^{-1})\}H(\underline{r};y) = e^{(\mu+2\nu)t} \sum_{k=-\infty}^{\infty} A_k y^k$$

and substitute for $H(\underline{r};y)$ from (2.6), we have

$$\sum_{k=-\infty}^{\infty} A_k y^k = e^{-(\mu+2\nu)t} \sum_{j=-\infty}^{\infty} I_j(2\nu t)y^j \sum_{i=-\infty}^{\infty} r_i y^i \quad . \tag{2.12}$$

Extracting the coefficient of y^k from (2.12) we get

$$A_k = e^{-(\mu+2\nu)t} \sum_{j=-\infty}^{\infty} r_{k-j} I_j(2\nu t)$$

and so

$$A_k = e^{-\mu t} - e^{-(\mu+2\nu)t} \sum_{j=-\infty}^{\infty} z_{k-j} I_j(2\nu t) \quad . \tag{2.13}$$

It now remains to evaluate the remaining equation in (2.4), namely

$$d(\log G)/dt = \lambda \sum_{k=-\infty}^{\infty} m_k(t)(z_k - 1) \quad . \tag{2.14}$$

On substituting for $m_k(t)$ and z_k from (1.3) and (2.11), respectively, we see that (2.14) becomes

$$d(\log G)/dt = -\lambda e^{\lambda t} \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} a_j A_i \sum_{k=-\infty}^{\infty} I_{k-j}(2\nu t) I_{k-i}(-2\nu t) \quad . \tag{2.15}$$

But

$$\sum_{k=-\infty}^{\infty} I_{k-j}(2\nu t) I_{k-i}(-2\nu t) = \sum_{k=-\infty}^{\infty} I_{j-k}(2\nu t) I_{k-i}(-2\nu t)$$

$$= \text{coeff. of } y^{j-i} \text{ in } \sum_{p=-\infty}^{\infty} I_p(2\nu t) y^p \sum_{q=-\infty}^{\infty} I_q(-2\nu t) y^q$$

which, from (2.9), is the same as

$$\text{coeff. of } y^{j-i} \text{ in } \exp\{\nu t(y+y^{-1})\}\exp\{-\nu t(y+y^{-1})\}$$

$$= \begin{cases} 1 & : \quad j=i \\ 0 & : \quad j \neq i \end{cases} \quad .$$

Hence the differential equation (2.15) becomes

$$d(\log G)/dt = -\lambda e^{\lambda t} \sum_{i=-\infty}^{\infty} a_i A_i$$

and this integrates to give

$$G = \text{const.} \exp\{-e^{\lambda t} \sum_{i=-\infty}^{\infty} a_i A_i\} \quad . \tag{2.16}$$

Thus the general solution to equation (2.2) is given by

$$G(\underline{z};t) = f(\ldots,A_i,\ldots)\exp\{-e^{\lambda t} \sum_{k=-\infty}^{\infty} a_k A_k\} \tag{2.17}$$

where f is an arbitrary function.

Expression (2.17) is the general integral of the Lagrange linear equation (2.2), and to derive the required integral we find that solution (2.17) which satisfies the initial condition (2.3). Now

$$I_j(0) = \begin{cases} 1 : j=0 \\ 0 : j\neq 0 \end{cases} , \tag{2.18}$$

and so when $t=0$ expression (2.11) gives

$$z_k = 1 - A_k \quad .$$

Hence we require

$$f(\ldots,A_i,\ldots) = \prod_{k=-\infty}^{\infty} (1-A_k)^{a_k} \quad . \tag{2.19}$$

We now replace the constants A_i $(-\infty < i < \infty)$ in the general solution (2.17) by expressions (2.13), namely

$$A_i = -Q_i(2\nu t) \qquad (-\infty < i < \infty) \tag{2.20}$$

where

$$Q_i(2\nu t) = e^{-(\mu+2\nu)t} \sum_{j=-\infty}^{\infty} z_{i-j} I_j(2\nu t) - e^{-\mu t} \quad . \tag{2.21}$$

This procedure yields the solution

$$G(\underline{z};t) = \prod_{i=-\infty}^{\infty} [g_i(\underline{z};t)]^{a_i} \tag{2.22}$$

where

$$g_i(\underline{z};t) = [1+Q_i(2\nu t)]\exp\{(e^{\lambda t}-1)Q_i(2\nu t)\} \tag{2.23}$$

is the probability generating function conditional on an initial population with a single member only in colony i. Without any loss of generality it is sufficient to consider $g_0(\underline{z};t)$. The solution (2.22)-(2.23) was verified by back-substitution into the

partial differential equation (2.2).

It follows from (2.18) and (2.21) that when $t=0$

$$Q_i(0) = z_i - 1 \quad ,$$

whence (2.23) becomes

$$g_i(\underline{z};0) = z_i \qquad (-\infty < i < \infty)$$

as required. Moreover, putting $y=1$ and $x=\nu t$ in (2.9) we get

$$\sum_{j=-\infty}^{\infty} I_j(2\nu t) = e^{2\nu t} \quad . \tag{2.24}$$

Thus when $\underline{z}=\underline{1}$ the function $Q_i(2\nu t)$ is identically zero and so $g_i(\underline{1};t) \equiv 1$, and it follows from (2.21) that when $|z_i| \leq 1$

$$0 \geq Q_i(2\nu t) \geq -2e^{-\mu t}$$

whence $|g_i(\underline{z};t)| \leq 1 \ (-\infty < i < \infty)$.

8.2.ii Approximate probabilities for a simple initial population

Let the initial population at time $t=0$ simply consist of a single individual in colony 0. We may evaluate the probabilities $p(\underline{x};t)$ by expanding $g_0(\underline{z};t)$ in terms of the dummy variables z_i and writing down the coefficient of $\prod_{j=-\infty}^{\infty} z_j^{x_j}$. This procedure gives

$$\begin{aligned} p(\underline{x};t) = {} & [(1-e^{-\mu t}) + n(e^{\lambda t}-1)^{-1}] \exp\{-e^{-\mu t}(e^{\lambda t}-1)\} \\ & \times (e^{\lambda t}-1)^n e^{-n(\mu+2\nu)t} \prod_{k=-\infty}^{\infty} (x_k!)^{-1} [I_k(2\nu t)]^{x_k} \quad , \end{aligned} \tag{2.25}$$

where the population vector is defined as $\underline{x} = (\ldots,x_k,\ldots)$ and

$$n = \sum_{j=-\infty}^{\infty} x_j$$

denotes the total population size at time t. In particular, the

'extinction' probability $p(\underline{0};t)$ is given by

$$p(\underline{0};t) = (1-e^{-\mu t})\exp\{-e^{-\mu t}(e^{\lambda t}-1)\} \tag{2.26}$$

which is identical with the 'extinction' probability (4.2.24) for the spatially homogeneous two-colony model. This is to be expected, because the total population size does not depend on how the population is partitioned into colonies. Note that the comments following expression (4.2.24) are equally applicable here.

Now for fixed i and large t

$$I_i(2\nu t) = (4\pi\nu t)^{-\frac{1}{2}}e^{2\nu t}\{1 - \frac{(4i^2-1)}{16\nu t} + O(t^{-2})\} \tag{2.27}$$

(Abramowitz and Stegun (1965), result (9.7.1)), and on using (2.27) we see that expression (2.25) simplifies to give

$$p(\underline{x};t) \sim (4\pi\nu t)^{-\frac{1}{2}n}\exp\{n(\lambda-\mu)t - e^{(\lambda-\mu)t}\} \prod_{k=-\infty}^{\infty}(x_k!)^{-1} \quad . \tag{2.28}$$

We may derive the marginal p.g.f. $g_0(z_k;t)$, for the number of individuals in colony k at time t, by placing $z_j=1$ $(-\infty < j < \infty;\ j\neq k)$ in (2.23) and then using result (2.24). This procedure gives

$$g_0(z_k;t) = [1 + e^{-(\mu+2\nu)t}(z_k-1)I_k(2\nu t)]\exp\{(e^{\lambda t}-1)e^{-(\mu+2\nu)t} \times (z_k-1)I_k(2\nu t)\} \quad . \tag{2.29}$$

Extracting the coefficient of $z_k^{x_k}$ in (2.29) we obtain the corresponding marginal probabilities

$$p(x_k;t) = [1 - e^{-(\mu+2\nu)t}I_k(2\nu t) + x_k(e^{\lambda t}-1)^{-1}] \times \exp\{-(e^{\lambda t}-1)e^{-(\mu+2\nu)t}I_k(2\nu t)\}[(e^{\lambda t}-1)e^{-(\mu+2\nu)t}I_k(2\nu t)]^{x_k}(x_k!)^{-1} \quad . \tag{2.30}$$

8.2.iii Asymptotic moments for a simple initial population

Moments of this process may be obtained by differentiation of $G(\underline{z};t)$ with respect to the z_i the appropriate number of times, and evaluation of the result at $\underline{z}=\underline{1}$. Evaluating the first-order moment $\bar{m}_i(t)$, we have

$$\bar{m}_i(t) = e^{(\lambda-\mu-2\nu)t} \sum_{j=-\infty}^{\infty} a_j I_{i-j}(2\nu t) \tag{2.31}$$

which agrees exactly with Bailey's expression (1.3) for $m_i(t)$. Thus the approximation leaves the means unaltered.

Let us now determine the variances and covariances $\bar{V}_{ij}(t)$ $(-\infty < i,j < \infty)$ for the simplified case when the initial population consists of a single individual at colony 0. Differentiating $g_0(\underline{z};t)$ twice with respect to the dummy variables z_k and placing $\underline{z}=\underline{1}$ we obtain

$$\begin{aligned} \bar{V}_{ij}(t) &= -e^{-2(\mu+2\nu)t} I_i(2\nu t) I_j(2\nu t) \qquad (i\neq j) \\ \bar{V}_{ii}(t) &= e^{(\lambda-\mu-2\nu)t} I_i(2\nu t) - e^{-2(\mu+2\nu)t} I_i^2(2\nu t) \quad . \end{aligned} \tag{2.32}$$

On using result (2.27) we have

$$\bar{V}_{ij}(t) = o(e^{-2\mu t}) \;\; ; \;\; \bar{V}_{ii}(t) = m_i(t) + o(e^{-2\mu t}) \qquad (i\neq j) \; . \tag{2.33}$$

Thus as $t\to\infty$ the covariances between the colonies approach zero, and within each colony the ratio $\bar{m}_i(t)$ to $\bar{V}_{ii}(t)$ approaches unity. This zero covariance is to be expected because of the dominating effect of the deterministic births. Bailey (1968) obtains an involved form for the generating function of the variances and covariances for the original unmodified process, and he shows that this form greatly simplifies when $\lambda=0$ to give expression (2.32). Note that in the general case $(\lambda\neq 0)$ we

expect the variance of the modified process to be less than that of the ordinary process because the births are deterministic.

Expressions (2.32) were verified by back-substitution into their own system of differential equations. We obtained this system by differentiating the differential equation (2.2) twice with respect to the z_i and then placing $\underline{z}=\underline{1}$.

When t is large we may simplify the first- and second-order moments by replacing the modified Bessel function by its asymptotic value. The substitution of (2.27) into (2.31) and (2.32) yields for $a_0=1$, $a_i=0$ $(i\neq 0)$ and $-\infty < i,j < \infty$

$$m_i(t) = e^{(\lambda-\mu)t}(4\pi\nu t)^{-\frac{1}{2}}\{1 - \frac{(4i^2-1)}{16\nu t} + O(t^{-2})\}$$

with

$$\overline{V}_{ij}(t) = -e^{-2\mu t}(4\pi\nu t)^{-1}\{1 - \frac{2(i^2+j^2)-1}{8\nu t} + O(t^{-2})\} \qquad (i\neq j)$$

and

$$\overline{V}_{ii}(t) = e^{(\lambda-\mu)t}(4\pi\nu t)^{-\frac{1}{2}}\{1 - \frac{(4i^2-1)}{16\nu t} + O(t^{-2})\}$$

$$- e^{-2\mu t}(4\pi\nu t)^{-1}\{1 - \frac{(4i^2-1)}{8\nu t} + O(t^{-2})\} .$$

8.3 The Effect of Immigration when $\lambda=0$

8.3.i Probability generating function

We shall now revert to the original unmodified process of section 8.1, and consider the effect of introducing immigration at rate α into colony 0 from outside the system. The equation for the probability generating function $G(\underline{z};t)$ corresponding to (1.2) is given by

$$(3.1)\quad \frac{\partial G}{\partial t} = \sum_{k=-\infty}^{\infty} \{(\lambda z_k-\mu)(z_k-1) + \nu(z_{k+1}-z_k)+\nu(z_{k-1}-z_k)\}\frac{\partial G}{\partial z_k} + \alpha(z_0-1)G$$

which has the associated auxiliary equations $(-\infty < k < \infty)$

$$(3.2) \qquad \frac{dt}{1} = \frac{-dz_k}{(\lambda z_k - \mu)(z_k - 1) + \nu(z_{k+1} - z_k) + \nu(z_{k-1} - z_k)} = \frac{dG}{\alpha(z_0 - 1)G} \quad .$$

In order to avoid the usual difficulty of the non-linearity in z_k, let us consider the special case of $\lambda = 0$. The first set of equations in (3.2) are now identical with the first set in (2.4). Hence expressions (2.11) and (2.13) hold for this model also; in particular

$$(3.3) \qquad z_0 - 1 = -e^{(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} A_i I_i(-2\nu t) \quad .$$

If we use (3.3) the last equation in (3.2) may now be written as

$$d(\log G)/dt = -\alpha e^{(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} A_i I_i(-2\nu t)$$

which integrates to give

$$(3.4) \qquad G = \text{const.} \exp\{-\alpha \sum_{i=-\infty}^{\infty} A_i P_i(t)\}$$

where

$$(3.5) \qquad P_i(t) = \int_0^t e^{(\mu+2\nu)\tau} I_i(-2\nu\tau)\, d\tau \quad .$$

Thus the general solution to equation (3.1) is given by

$$(3.6) \qquad G(\underline{z};t) = f(\ldots, A_i, \ldots) \exp\{-\alpha \sum_{j=-\infty}^{\infty} A_j P_j(t)\}$$

where f is an arbitrary function.

Expression (3.6) is the general integral of the Lagrange linear equation (3.1), and to derive the required integral we find that solution (3.6) which satisfies the initial condition

$$(3.7) \qquad G(\underline{z};0) = \prod_{k=-\infty}^{\infty} z_k^{a_k} \quad .$$

Now we have previously shown that when $t=0$

$$z_k = 1-A_k \quad ,$$

and so we require

$$f(\ldots,A_i,\ldots) = \prod_{k=-\infty}^{\infty} (1-A_k)^{a_k} \quad .$$

We now replace the constants A_i $(-\infty < i < \infty)$ in the general solution (3.6) by expressions (2.20), namely

$$A_i = -Q_i(2\nu t) \qquad (-\infty < i < \infty) \; .$$

This procedure yields the solution

$$(3.8) \qquad G(\underline{z};t) = \exp\{\alpha \sum_{j=-\infty}^{\infty} Q_j(2\nu t)P_j(t)\} \prod_{k=-\infty}^{\infty} [1+Q_k(2\nu t)]^{a_k} \quad .$$

Note that if $z_i=z$ $(-\infty < i < \infty)$ in (2.21) then result (2.24) yields

$$Q_i(2\nu t) = e^{-\mu t}(z-1) \qquad (-\infty < i < \infty) \; .$$

Hence the p.g.f. of the total population size is given by

$$(3.9) \qquad G(z;t) = [1+e^{-\mu t}(z-1)]^{\Sigma a_i} \exp\{(\alpha/\mu)(z-1)(1-e^{-\mu t})\} \quad ,$$

and expression (3.9) is the well-known result for the simple immigration-death process (Cox and Miller (1965), p.168).

Suppose that immigration is not present $(\alpha=0)$ and the initial population at time $t=0$ is given by

$$a_0=1, \; a_i=0 \; (i\neq 0) \; .$$

Then on noting that the function $Q_i(2\nu t)$ is defined by expression (2.21), we see from either (2.23) with $\lambda=0$, or alternatively (3.8), that

$$(3.10) \qquad G(\underline{z};t) = 1 + e^{-(\mu+2\nu)t} \sum_{j=-\infty}^{\infty} z_j I_j(2\nu t) - e^{-\mu t} \quad .$$

The original individual is unable to multiply as births are suppressed, and at any particular time $t \geq 0$ it must be either situated in one of the colonies or be dead. It follows from (3.10) that the probabilities of these events are

$$\Pr[\text{dead by time } t] = 1 - e^{-\mu t}$$

$$\Pr[\text{alive and in colony } i \text{ at time } t] = e^{-(\mu+2\nu)t} I_i(2\nu t) \quad .$$

8.3.ii An alternative form of solution

Although (3.8) is the solution to equation (3.1), it is possible to simplify it by replacing the term

$$R(\underline{z};t) \equiv \sum_{i=-\infty}^{\infty} P_i(t) Q_i(2\nu t)$$

with an alternative expression. Substituting for $Q_i(2\nu t)$ from (2.21) we have

$$(3.11) \qquad R(\underline{z};t) = \sum_{i=-\infty}^{\infty} P_i(t)\left[e^{-(\mu+2\nu)t} \sum_{j=-\infty}^{\infty} z_{i-j} I_j(2\nu t) - e^{-\mu t}\right] \quad ,$$

whilst from (3.5) and (2.9)

$$\sum_{i=-\infty}^{\infty} P_i(t) y^i = \sum_{i=-\infty}^{\infty} \int_0^t e^{(\mu+2\nu)s} I_i(-2\nu s) y^i ds$$

$$= \int_0^t \exp\{s[(\mu+2\nu) - \nu(y+y^{-1})]\} ds$$

$$(3.12) \qquad = [(\mu+2\nu) - \nu(y+y^{-1})]^{-1} [\exp\{t[(\mu+2\nu) - \nu(y+y^{-1})]\} - 1] \quad .$$

Placing $y=1$ in (3.12) we get

$$(3.13) \qquad \sum_{i=-\infty}^{\infty} P_i(t) = \mu^{-1}(e^{\mu t} - 1)$$

and so (3.11) becomes

$$R(\underline{z};t) = e^{-(\mu+2\nu)t} \sum_{h=-\infty}^{\infty} z_h \sum_{i=-\infty}^{\infty} P_i(t) I_{i-h}(2\nu t) - \mu^{-1}(1-e^{-\mu t}) \quad .$$

(3.14)

But

$$e^{-(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} P_i(t) I_{i-h}(2\nu t)$$

is the coefficient of y^h in

$$e^{-(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} P_i(t) y^i \sum_{j=-\infty}^{\infty} I_j(2\nu t) y^{-j} \quad ,$$

that is, if we use (2.9) and (3.12), in

$$(3.15) \qquad [(\mu+2\nu)-\nu(y+y^{-1})]^{-1}[1 - \exp\{-t[(\mu+2\nu)-\nu(y+y^{-1})]\}] \quad .$$

LEMMA : The identity

$$(3.16) \qquad [1-\theta(z+z^{-1})]^{-1} \equiv (1-4\theta^2)^{-\frac{1}{2}} \sum_{n=-\infty}^{\infty} \beta^{|n|} z^n \qquad (0 < \theta < \tfrac{1}{2})$$

holds, where

$$(3.17) \qquad \beta,\beta' = [1 \mp (1-4\theta^2)^{\frac{1}{2}}]/(2\theta)$$

respectively, and $\beta<|z|<\beta'$.

Proof of lemma : The result follows if we write

$$[1-\theta(z+z^{-1})]^{-1} = \frac{1}{\theta(\beta'-\beta)}\left[1 + \frac{z/\beta'}{1-z/\beta'} + \frac{\beta/z}{1-\beta/z}\right]$$

and expand the right hand side in powers of z .

Thus defining $\theta=\nu/(\mu+2\nu)$, and noting that $0 < \theta < \frac{1}{2}$ for $\mu > 0$, we see that (3.15) may be written as

$$(\mu+2\nu)^{-1}(1-4\theta^2)^{-\frac{1}{2}}\left[\sum_{n=-\infty}^{\infty} \beta^{|n|} y^n\right]\left[1 - e^{-(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} I_i(2\nu t) y^i\right] \quad .$$

Hence the coefficient of y^h in expression (3.15) is given by

$$(3.18) \qquad (\mu+2\nu)^{-1}(1-4\theta^2)^{-\frac{1}{2}}\{\beta^{|h|} - e^{-(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} \beta^{|h-i|} I_i(2\nu t)\} \quad .$$

Comparing (3.8), (3.14) and (3.18) we obtain the alternative solution

$$G(\underline{z};t) = \exp\{-(\alpha/\mu)(1-e^{-\mu t})\}\exp\{(\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}}$$

$$\times \sum_{h=-\infty}^{\infty} z_h[\beta^{|h|} - e^{-(\mu+2\nu)t} \sum_{i=-\infty}^{\infty} \beta^{|h-i|} I_i(2\nu t)]\} \prod_{i=-\infty}^{\infty} [1+Q_i(2\nu t)]^{a_i}. \tag{3.19}$$

Expression (3.19) was verified by back-substitution into equation (3.1). Note that

$$G(\underline{z};0) = \prod_{i=-\infty}^{\infty} z_i^{a_i} ,$$

$G(\underline{1};t)\equiv 1$ and $|G(\underline{z};t)| \leq 1$ for $|z_i| \leq 1$ $(-\infty < i < \infty)$ as required.

If all the colonies are initally empty, so that $a_i=0$ $(-\infty < i < \infty)$, $G(\underline{z};t)$ may clearly be factored in the form

$$G(\underline{z};t) = \exp\{-(\alpha/\mu)(1-e^{-\mu t})\} \prod_{i=-\infty}^{\infty} \exp\{(\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}} z_i f_i(t)\} \tag{3.20}$$

where

$$f_i(t) = \beta^{|i|} - e^{-(\mu+2\nu)t} \sum_{n=-\infty}^{\infty} \beta^{|i-n|} I_n(2\nu t) \quad . \tag{3.21}$$

The probabilities $p(\underline{x};t)$ may then be written as

$$p(\underline{x};t) = \exp\{-(\alpha/\mu)(1-e^{-\mu t})\} \prod_{n=-\infty}^{\infty} (x_n!)^{-1}[(\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}} f_n(t)]^{x_n} \tag{3.22}$$

where the vector $\underline{x}=(\ldots,x_n,\ldots)$ for $x_n=0,1,2,\ldots$ $(-\infty < n < \infty)$.

8.3.iii Moments

We may evaluate moments of the process by differentiating the p.g.f. (3.19) with respect to the dummy variables z_i the appropriate number of times and placing $\underline{z}=\underline{1}$. In particular, the means $m_i(t)$ are given by

$$m_i(t) = (\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}} f_i(t) + e^{-(\mu+2\nu)t} \sum_{n=-\infty}^{\infty} a_n I_{n-i}(2\nu t) \tag{3.23}$$

whilst the variances and covariances $V_{ij}(t)$ are given by

$$(3.24)\quad V_{ij}(t) = m_i(t)\delta^i_j - e^{-2(\mu+2\nu)t}\sum_{n=-\infty}^{\infty} a_n I_{n-i}(2\nu t) I_{n-j}(2\nu t) .$$

Here

$$(3.25)\quad \delta^i_j := \begin{cases} 1 : j=i \\ 0 : j\neq i \end{cases}$$

denotes the Kronecker delta function. These results were verified by reference to their own system of differential equations. This system is given in the following section by equations (4.1) and (4.18) with $\lambda=0$.

Now expression (3.23) may be written in the form

$$(3.26)\quad m_i(t) = (\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}}\beta^{|i|} + e^{-(\mu+2\nu)t}\sum_{n=-\infty}^{\infty}[a_n-(\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}}\beta^{|n|}]I_{n-i}(2\nu t) .$$

Thus, if the initial population is geometrically distributed amongst the colonies with

$$(3.27)\quad a_i = (\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}}\beta^{|i|} \qquad (-\infty < i < \infty) ,$$

we have

$$m_i(t) \equiv a_i \qquad (-\infty < i < \infty)$$

for all $t \geq 0$ and in this sense we may call an initial population distribution $\{a_i\}$ which satisfies (3.27) 'stationary'. Note that for large t , application of the asymptotic result (2.27) to (3.26) yields

$$(3.28)\quad m_i(t) \sim \beta^{|i|}(\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}} + e^{-\mu t}(4\pi\nu t)^{-\frac{1}{2}}[\sum_{j=-\infty}^{\infty} a_j - \alpha/\mu] .$$

Thus as $t\to\infty$ the $m_i(t)$ approach their limiting values

$$m_i(\infty) = \beta^{|i|}(\alpha\theta/\nu)(1-4\theta^2)^{-\frac{1}{2}}$$

at rate

$$t^{-\frac{1}{2}}e^{-\mu t} \quad ,$$

unless of course $\Sigma a_j = \alpha/\mu$ when the approach to the limits is even faster.

8.3.iv Equilibrium distribution

If the death rate μ is positive we clearly expect an equilibrium situation to develop as t tends to infinity, because births are suppressed. Now we see from (2.24) that for any constants c_i such that $|c_i| \leq 1$ $(-\infty < i < \infty)$

$$(3.29) \qquad \Big| \sum_{i=-\infty}^{\infty} c_i I_i(2\nu t) \Big| \leq \sum_{i=-\infty}^{\infty} I_i(2\nu t) = e^{2\nu t} \quad ,$$

because $I_i(2\nu t)$ is real and non-negative when i is an integer and $t \geq 0$. Hence if we let $t \to \infty$ in (3.19) it follows that for $\mu > 0$ the p.g.f. $G(\underline{z};t)$ has a limiting value given by

$$(3.30) \qquad G(\underline{z};\infty) = e^{-\alpha/\mu} \exp\{\alpha(\mu^2+4\mu\nu)^{-\frac{1}{2}} \sum_{n=-\infty}^{\infty} z_n \beta^{|n|}\} \quad .$$

An equilibrium situation therefore exists with associated probabilities

$$(3.31) \qquad p(\underline{x};\infty) = e^{-\alpha/\mu} \prod_{n=-\infty}^{\infty} (x_n!)^{-1} [\alpha(\mu^2+4\mu\nu)^{-\frac{1}{2}} \beta^{|n|}]^{x_n} \quad .$$

The equilibrium p.g.f. (3.30) may be written as the product

$$(3.32) \qquad G(\underline{z};\infty) = e^{-\alpha/\mu} \prod_{i=-\infty}^{\infty} \exp\{\alpha(\mu^2+4\mu\nu)^{-\frac{1}{2}} z_i \beta^{|i|}\} \quad ,$$

and so in equilibrium each colony behaves as though it were independent of all other colonies, the size of colony i being distributed as a Poisson variable with parameter

$$\alpha(\mu^2+4\mu\nu)^{-\frac{1}{2}} \beta^{|i|} \quad .$$

Note that from (3.17) we have $0 < \beta < 1$ and so both the means and the variances decrease geometrically away from the origin at rate β.

8.4 The Effect of Immigration when $\lambda \geq 0$; First- and Second-Order Moments

As the nonlinearity of equation (3.1) when $\lambda > 0$ makes it extremely difficult to solve, we obtained previously the p.g.f $G(\underline{z};t)$ when $\lambda=0$ and from it we derived the first- and second-order moments. We shall now obtain expressions for moments of the first- and second-order which are valid for all $\lambda \geq 0$.

Let $m_i(t)$ denote the mean number of individuals in colony i at time t . Then differentiating (3.1) with respect to z_i and placing $\underline{z=1}$ we obtain

$$dm_i(t)/dt = (\lambda-\mu-2\nu)m_i(t) + \nu m_{i-1}(t) + \nu m_{i+1}(t) + \alpha\delta_i^0 \quad (-\infty < i < \infty) \tag{4.1}$$

where

$$m_i(0) = a_i \tag{4.2}$$

and δ_i^0 is the Kronecker delta function (3.25). Put $\xi=\lambda-\mu-2\nu$ and denote the Laplace transform (2.3.3) of a function $g(t)$ by

$$\mathcal{L}[g(t)] \equiv g^*(s) = \int_0^\infty e^{-st} g(t)dt \quad , \quad \Re(s) > 0 . \tag{4.3}$$

Applying the transformation (4.3) to equations (4.1) we get the set of simultaneous equations

$$sm_i^*(s) - a_i = \xi m_i^*(s) + \nu m_{i-1}^*(s) + \nu m_{i+1}^*(s) + \alpha s^{-1}\delta_i^0 \quad (-\infty < i < \infty) , \tag{4.4}$$

for $\Re(s)$ large enough to ensure the existence of the $m_i^*(s)$. Introduce the generating function

$$W(z;t) = \sum_{i=-\infty}^{\infty} m_i(t) z^i$$

with Laplace transform

$$W^*(z;s) = \sum_{i=-\infty}^{\infty} m_i^*(s) z^i \quad .$$

Then multiplying both sides of (4.4) by z^i and summing over the range $-\infty < i < \infty$ we obtain

$$sW^*(z;s) - W(z;0) = (\nu z+\xi+\nu z^{-1})W^*(z;s) + \alpha s^{-1}$$

which, on rearrangement, yields

$$W^*(z;s) = -[W(z;0)+\alpha s^{-1}][\nu z+(\xi-s)+\nu z^{-1}]^{-1} \quad . \tag{4.5}$$

Now the equation

$$\nu z^2+(\xi-s)z+\nu = 0$$

has roots

$$z_1, z_2 = [(s-\xi) \pm \{(s-\xi)^2-4\nu^2\}^{\frac{1}{2}}]/(2\nu) \quad , \tag{4.6}$$

and so

$$W^*(z;s) = -z[W(z;0)+\alpha s^{-1}][\nu(z-z_1)(z-z_2)]^{-1} \quad . \tag{4.7}$$

If we split (4.7) into partial fractions we obtain

$$W^*(z;s) = (1/\nu)[W(z;0)+\alpha s^{-1}](z_1-z_2)^{-1}[(1-z/z_1)^{-1} + (z_2/z)(1-z_2/z)^{-1}] \quad . \tag{4.8}$$

Thus if s and z are such that $|z_2| < |z| < |z_1|$, then expansion of (4.8) in powers of z yields

$$W^*(z;s) = [W(z;0)+\alpha s^{-1}][(s-\xi)^2-4\nu^2]^{-\frac{1}{2}} \sum_{j=-\infty}^{\infty} z_2^{|j|} z^j \quad . \tag{4.9}$$

Inversion of this function proceeds as follows. From Abramowitz and Stegun (1965), result (29.3.59),

$$\mathcal{L}[(2\nu)^i I_i(2\nu t)] = [s-(s^2-4\nu^2)^{\frac{1}{2}}]^i (s^2-4\nu^2)^{-\frac{1}{2}} \quad ,$$

and so using Abramowitz and Stegun (1965), result (29.2.12), we have

$$\mathcal{L}[e^{\xi t}(2\nu)^i I_i(2\nu t)] = [(s-\xi)-\{(s-\xi)^2-4\nu^2\}^{\frac{1}{2}}]^i[(s-\xi)^2-4\nu^2]^{-\frac{1}{2}}$$
$$= z_2^i(2\nu)^i[(s-\xi)^2-4\nu^2]^{-\frac{1}{2}} \quad . \tag{4.10}$$

Also

$$\mathcal{L}[1] = s^{-1} \quad . \tag{4.11}$$

Hence, if we use the convolution theorem (Abramowitz and Stegun (1965), result (29.2.8)), which states that for general functions $f_1(t)$ and $f_2(t)$

$$\mathcal{L}[\int_o^t f_1(t-\tau)f_2(\tau)d\tau] = \mathcal{L}[f_1(t)]\mathcal{L}[f_2(t)] \quad , \tag{4.12}$$

a combination of (4.10) and (4.11) will give

$$\mathcal{L}[\int_o^t e^{\xi\tau} I_i(2\nu\tau)d\tau] = z_2^{|j|}[s\{(s-\xi)^2-4\nu^2\}^{\frac{1}{2}}]^{-1} \quad . \tag{4.13}$$

So a comparison of (4.9), (4.10) and (4.13) yields

$$W(z;t) = e^{\xi t}\sum_{n=-\infty}^{\infty} a_n z^n \sum_{j=-\infty}^{\infty} I_j(2\nu t)z^j + \alpha\int_o^t e^{\xi\tau}\sum_{j=-\infty}^{\infty} z^j I_j(2\nu\tau)d\tau \quad . \tag{4.14}$$

Comparing coefficients of z^i on both sides of (4.14) we obtain the solution

$$m_i(t) = e^{(\lambda-\mu-2\nu)t}\sum_{n=-\infty}^{\infty} a_n I_{i-n}(2\nu t) + \alpha\int_o^t e^{(\lambda-\mu-2\nu)\tau} I_i(2\nu\tau)d\tau \quad . \tag{4.15}$$

Expression (4.15) was verified by back-substitution into equation (4.1). It clearly splits into two components. The first is identical with Bailey's result (1.3) for the model without immigration. The second is the contribution to the means from immigration. It may be shown that when $\lambda=0$ expression (4.15) is identical with (3.23).

The variances and covariances $V_{ij}(t)$ $(-\infty < i,j < \infty)$ may be derived by the use of a similar technique. As it is simpler to work with the cumulant generating function than with the probability generating function $G(\underline{z};t)$, let us replace the dummy variables z_k in (3.1) by $\exp(\theta_k)$ $(-\infty < k < \infty)$ and write

$$K(\underline{\theta};t) = \log G(\underline{z};t)$$

where the infinite-dimensional vector $\underline{\theta} = (\ldots,\theta_k,\ldots)$ has θ_k in the k^{th} position. Equation (3.1) may now be written as

$$\frac{\partial K}{\partial t} = \sum_{k=-\infty}^{\infty} [\lambda(e^{\theta_k}-1)+\mu(e^{-\theta_k}-1)+\nu(e^{\theta_{k+1}-\theta_k}-1)+\nu(e^{\theta_{k-1}-\theta_k}-1)]\frac{\partial K}{\partial \theta_k} + \alpha(e^{\theta_O}-1) \quad . \tag{4.16}$$

Using (1.1), we see that

$$K(\underline{\theta};O) = \sum_{i=-\infty}^{\infty} a_i\theta_i \quad . \tag{4.17}$$

To derive the equations for the variances and covariances we first differentiate both sides of (4.16) with respect to θ_i and θ_j and then place $\underline{\theta} = \underline{O}$. Care must be taken because each $\exp(\theta_k)$ occurs in three successive terms of the summation. Using this procedure we get the equations $(-\infty < i,j < \infty)$

$$\begin{aligned} dV_{ij}(t)/dt = {} & 2(\lambda-\mu-2\nu)V_{ij}(t) + \nu[V_{i,j-1}(t)+V_{i,j+1}(t)+V_{i-1,j}(t) \\ & +V_{i+1,j}(t)] + \alpha\delta_i^O\delta_j^O - \nu[m_i(t)+m_{i+1}(t)]\delta_j^{i+1} \\ & - \nu[m_i(t)+m_{i-1}(t)]\delta_j^{i-1} + [(\lambda+\mu+2\nu)m_i(t)+\nu m_{i-1}(t) \\ & +\nu m_{i+1}(t)]\delta_i^j \end{aligned} \tag{4.18}$$

where the $m_i(t)$ are defined by expression (4.15) and δ_j^i represents the Kronecker delta function (3.25).

As all the variances and covariances are initially zero, on taking the Laplace transform of both sides of (4.18) we obtain

$$
\begin{aligned}
sV^*_{ij}(s) &= 2\xi V^*_{ij}(s) + \nu(V^*_{i,j-1}(s)+V^*_{i,j+1}(s)+V^*_{i-1,j}(s)+V^*_{i+1,j}(s)) \\
&+ \alpha s^{-1}\delta^0_i\delta^0_j - \nu(m^*_i(s)+m^*_{i+1}(s))\delta^{i+1}_j - \nu(m^*_i(s)+m^*_{i-1}(s))\delta^{i-1}_j \\
&+ [(\lambda+\mu+2\nu)m^*_i(s)+\nu m^*_{i-1}(s)+\nu m^*_{i+1}(s)]\delta^i_j
\end{aligned} \tag{4.19}
$$

where $\xi=\lambda-\mu-2\nu$. Introduce the generating function

$$K(x,y;t) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} V_{ij}(t)x^i y^j$$

with Laplace transform

$$K^*(x,y;s) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} V^*_{ij}(s)x^i y^j \quad .$$

Then multiplication of both sides of (4.19) by $x^i y^j$ and summation of the result over the range $-\infty < i,j < \infty$ gives

$$
\begin{aligned}
sK^*(x,y;s) &= [2\xi+\nu(x+x^{-1}+y+y^{-1})]K^*(x,y;s) + \alpha s^{-1} \\
&- \nu(x+x^{-1}+y+y^{-1})W^*(xy;s) + [(\lambda+\mu+2\nu)+\nu(xy)+\nu(xy)^{-1}]W^*(xy;s)
\end{aligned}
$$

which, on rearrangement, yields

$$
\begin{aligned}
K^*(x,y;s) &= [(s-2\xi)-\nu(x+x^{-1}+y+y^{-1})]^{-1}[(\lambda+\mu+2\nu) \\
&\qquad -\nu(x+x^{-1}+y+y^{-1}-xy-x^{-1}y^{-1})]W^*(xy;s) \\
&+ [(s-2\xi)-\nu(x+x^{-1}+y+y^{-1})]^{-1}\alpha s^{-1}
\end{aligned} \tag{4.20}
$$

for

$$\Re(s) > \max\{0,\ 2\xi+\nu(x+x^{-1}+y+y^{-1})\} \quad .$$

Bailey (1968) considers the system of differential equations for the variances and covariances when immigration is suppressed, and he obtains a complicated expression for the generating function

$$V(x,y;t) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} V_{ij}(t)x^i y^j \quad .$$

Note that when $\alpha=0$ expressions (4.18) and (4.20) are identical with Bailey's expressions (26)-(28) and (30), respectively. He does not, however, obtain the general term $V_{ij}(t)$ from the generating function, but states that this appears to be a matter of some complexity involving the evaluation of a complex integral for the relevant term in the two-dimensional Laurent expansion of $V(x,y;t)$. We shall now obtain the variances and covariances $V_{ij}(t)$ by inverting (4.20) in a different manner to that of Bailey.

The expression

$$L^*(x,y;s) \equiv [\,(s-2\xi)-\nu(x+x^{-1}+y+y^{-1})\,]^{-1}$$

in (4.20) inverts immediately to yield

$$L(x,y;t) = \exp\{[\,2\xi+\nu(x+x^{-1})+\nu(y+y^{-1})\,]\,t\} \quad ,$$

which from (2.9) gives

$$L(x,y;t) = e^{2\xi t} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I_i(2\nu t) I_j(2\nu t) x^i y^j \quad . \tag{4.21}$$

Hence, on inverting (4.20) by means of the convolution theorem (4.12), we have

$$K(x,y;t) = \int_0^t L(x,y;t-\tau)\{[\,(\lambda+\mu+2\nu)-\nu(x+x^{-1}+y+y^{-1}-xy-x^{-1}y^{-1})\,]W(xy;\tau) + \alpha\}d\tau \quad ,$$

and extraction of the appropriate coefficients of $x^i y^j$ on both sides of this expression yields the solution $(-\infty < i,j < \infty)$

$$\begin{aligned} V_{ij}(t) = \int_0^t e^{2\xi(t-s)} \sum_{r=-\infty}^{\infty} [\,&(\lambda+\mu+2\nu) I_{i-r}(2\nu(t-s)) I_{j-r}(2\nu(t-s)) \\ &- \nu I_{i-1-r}(2\nu(t-s)) I_{j-r}(2\nu(t-s)) - \nu I_{i+1-r}(2\nu(t-s)) I_{j-r}(2\nu(t-s)) \\ &- \nu I_{i-r}(2\nu(t-s)) I_{j-1-r}(2\nu(t-s)) - \nu I_{i-r}(2\nu(t-s)) I_{j+1-r}(2\nu(t-s)) \end{aligned}$$

$$+ \nu I_{i-1-r}(2\nu(t-s))I_{j-1-r}(2\nu(t-s)) + \nu I_{i+1-r}(2\nu(t-s))I_{j+1-r}(2\nu(t-s))]$$

$$(4.22) \qquad \times \{e^{\xi s} \sum_{n=-\infty}^{\infty} a_n I_{r-n}(2\nu s) + \alpha \int_0^s e^{\xi\tau} I_r(2\nu\tau)d\tau\}ds$$

$$+ \alpha \int_0^t e^{2\xi(t-s)} I_i(2\nu(t-s))I_j(2\nu(t-s))ds \quad .$$

To verify expression (4.22) we first note that $V_{ij}(0) = 0$ $(-\infty < i,j < \infty)$ as required. Now

$$I_i(0) = \begin{cases} 1 : i=0 \\ 0 : i\neq 0 \end{cases} ,$$

whilst from Abramowitz and Stegun (1965), result (9.6.26),

$$\frac{d}{dt} I_i(2\nu t) = \nu[I_{i-1}(2\nu t) + I_{i+1}(2\nu t)] \quad .$$

Differentiation of (4.22) with respect to t and use of the above results retrieves the original differential equations (4.18).

8.5 The Negative Binomial Approximation

The comparison between the negative binomial approximation and the modified process was investigated at some length for the two-colony model in sections 4.3 and 7.3. On almost all accounts the former seemed to be the better approximation to the true distribution. We shall now derive an equivalent negative binomial approximation for the stepping-stone model, but we shall not include any analytic comparison between this distribution and the corresponding expression developed in section 8.2. Suffice it to say that the results for the two-colony comparison may well apply here.

Consider the p.g.f.

$$(5.1) \qquad G(\underline{z};t) = \prod_{k=-\infty}^{\infty} \left[\frac{1-c_k(t)z_k}{1-c_k(t)} \right]^{d_k(t)}$$

where $c_k(t)$ and $d_k(t)$ are unknown functions of time. On writing the infinite-dimensional vector $\underline{\theta} = (\ldots,\theta_k,\ldots)$, where $z_k=\exp(\theta_k)$ $(-\infty < k < \infty)$, we see that the c.g.f.

$$K(\underline{\theta};t) = \log G(\underline{z};t)$$

is given by the expression

$$(5.2) \qquad K(\underline{\theta};t) = \sum_{k=-\infty}^{\infty} d_k[\log(1-c_k e^{\theta_k}) - \log(1-c_k)] \quad .$$

Differentiation of (5.2) with respect to the θ_k gives

$$\frac{\partial K}{\partial \theta_k} = -d_k c_k e^{\theta_k}/(1-c_k e^{\theta_k})$$

$$\frac{\partial^2 K}{\partial \theta_k^2} = -d_k c_k/(1-c_k e^{\theta_k})^2$$

and

$$\frac{\partial^2 K}{\partial \theta_k \partial \theta_j} = 0 \qquad (k \neq j) \quad ,$$

whence placing $\theta_k=0$ $(-\infty < k < \infty)$ in these differential coefficients we obtain

$$(5.3) \qquad m_k(t) = -d_k c_k/(1-c_k)$$

together with

$$(5.4) \quad V_{kk}(t) = -d_k c_k/(1-c_k)^2 \quad ; \quad V_{kh}(t) \equiv 0 \qquad (-\infty < k,h < \infty;\ h \neq k) \quad .$$

Here we have equated the approximate means and variances derived from the p.g.f. (5.1) to the true means and variances, these being given in greatest generality by expressions (4.15) and (4.22), respectively. Solving for c_k and d_k from (5.3) and (5.4)

we get

$$c_k(t) = 1 - m_k(t)/V_{kk}(t)$$

(5.5)

$$d_k(t) = -m_k^2(t)/(V_{kk}(t)-m_k(t)) \quad .$$

It follows from (4.3.13) that if an equilibrium situation exists, expression (5.1) represents the product of the p.g.f.'s of an infinite number of independent simple birth-death-immigration processes. The birth, death and immigration parameters of the k^{th} process $(-\infty < k < \infty)$, denoted by λ_k', μ_k' and α_k' respectively, satisfy the relations

$$\lambda_k'/\mu_k' = c_k(\infty) \quad ; \quad \alpha_k'/\lambda_k' = -d_k(\infty) \qquad (-\infty < k < \infty) \quad .$$

CHAPTER 9

VELOCITY OF PROPAGATION FOR STEPPING-STONE MODELS

9.1 Introduction

Fisher (1936) and Kolmogorov, Petrovsky and Piscounov (1937) (henceforth abbreviated to KPP) independently attempted to determine the velocity of the wave of advance of an advantageous gene. For this purpose a population is assumed to be distributed in a linear habitat, such as a shore line. If at any point on the line a mutation occurs which confers a selective advantage, we may expect the mutants to diffuse into the surrounding population and eventually to take over the entire territory. The differential equation describing the velocity at which the advantageous gene spreads is a combination of a heat equation and a nonlinear term, and the authors mentioned above prove that this equation cannot admit velocities below a certain minimum c_m. Moreover, they show that all velocities in the range $[c_m,\infty)$ are possible. The Russian authors go on to prove that if the region occupied by an initial population of mutants is bounded, then two 'waves' travelling in opposite directions from that region will occur. Asymptotically the waves will have an invariant profile and a velocity equal to c_m.

Although no simple closed form solution to the original differential equation has yet been obtained, Montroll (1967) adds a further term to the equation which enables such a solution to be derived. There is controversy about the relevance of his solution to the original equation, but the paper certainly gives a considerable qualitative insight into the problem. The ideas of

Fisher and KPP are developed further by Canosa (1973). He proves that the propagation velocity of the waves is linearly proportional to their thickness and he derives an analytic solution for the travelling wave profiles. Canosa makes the interesting point that Fisher and KPP's equation is equivalent to a model used to describe the evolution of the neutron population in a nuclear reactor. The neutrons develop by the same physical processes as the mutants, that is, by diffusion and nonlinear local multiplication. Other interesting models for which similar results might reasonably be expected to hold include those developed by Marris (1970) for the spread of consumer demand and Zeldovitch (1951) for flame propagation.

Recent work in this field has centred around the spatial spread of epidemics. Kendall (1965) employs the techniques of Fisher and KPP to examine travelling epidemic waves; his results are achieved by an approximation of the spatial dependence with the introduction of a diffusion term in the equation of propagation. Mollison (1972a) shows that, when the spatial distribution of contacts is more spread than that described by the negative exponential distribution, no finite bound can be set on the velocity of propagation. Mollison (1972b) further suggests that the diffusion-term approximations may be reasonable if and only if the spatial distribution of contacts is less spread than that described by the negative exponential distribution.

In this Chapter I develop the velocity of propagation and the wave profile for a migrating population which spreads according to the stepping-stone model of Chapter 8. Unlike the authors above,

I shall not use a diffusion approximation to obtain the desired results (sections 9.2-9.6). Indeed, the use of such an approximation gives rise to a totally different expression for the velocity (section 9.7).

9.2 First-Order Moments with General Linear Migration

So far we have allowed migration to occur between nearest neighbours only. We shall now consider the more general migration distribution given by

(2.1) Pr[an individual in colony j migrates to colony $i+j$ in $(t,t+\delta t)$]

$$= \nu h_i \delta t + o(\delta t) \qquad (-\infty < i,j < \infty)$$

where

(2.2) $$\sum_{i=-\infty}^{\infty} h_i = 1 ; \qquad h_0 = 0 \quad .$$

Thus the total migration rate out of a colony is equal to ν. For this model the Kolmogorov forward differential equation for the p.g.f. $G(\underline{z};t)$ is given by

(2.3) $$\frac{\partial G}{\partial t} = \sum_{i=-\infty}^{\infty} \{\lambda(z_i^2 - z_i) + \mu(1-z_i) + \nu \sum_{j=-\infty}^{\infty} h_j (z_{i+j} - z_i)\} \frac{\partial G}{\partial z_i} \quad .$$

We may now obtain the system of difference-differential equations for $m_i(t)$, the mean number of individuals in colony i at time t, by differentiating both sides of equation (2.3) with respect to the dummy variables z_i and then placing $z_i = 1$ $(-\infty < i < \infty)$. This procedure yields

(2.4) $$dm_i(t)/dt = (\lambda - \mu - \nu) m_i(t) + \nu \sum_{j=-\infty}^{\infty} m_{i-j}(t) h_j \qquad (-\infty < i < \infty).$$

Note that we have suppressed immigration. On applying the Laplace

transformation (2.3.3) to both sides of equation (2.4) we have

$$(2.5) \qquad sm_i^*(s) - a_i = (\lambda-\mu-\nu)m_i^*(s) + \nu \sum_{j=-\infty}^{\infty} m_{i-j}^*(s)h_j$$

where

$$m_i(0) = a_i$$

denotes the initial number of individuals in colony i at time $t=0$. Define the generating functions

$$M^*(s,z) = \sum_{i=-\infty}^{\infty} m_i^*(s)z^i \quad \text{and} \quad H(z) = \sum_{i=-\infty}^{\infty} h_i z^i$$

where $|z| < 1$. Then multiplication of both sides of (2.5) by z^i and summation of the result over the range $-\infty < i < \infty$ yields

$$M^*(s,z)[s-\lambda+\mu+\nu-\nu H(z)] = \sum_{i=-\infty}^{\infty} a_i z^i ,$$

and this expression inverts to give

$$\sum_{i=-\infty}^{\infty} m_i(t)z^i = e^{(\lambda-\mu-\nu)t}\exp\{\nu t \sum_{j=-\infty}^{\infty} h_j z^j\}[\sum_{k=-\infty}^{\infty} a_k z^k] \quad .$$

In particular, if the initial population at time $t=0$ consists of a single individual in colony 0, so that $a_0=1$ and $a_i=0$ $(i\neq 0)$, we have

$$(2.6) \qquad \sum_{i=-\infty}^{\infty} m_i(t)z^i = e^{(\lambda-\mu-\nu)t}\exp\{\nu t \sum_{j=-\infty}^{\infty} h_j z^j\} \quad .$$

9.3 Velocities of Propagation for Nearest-Neighbour Migration

9.3.i Approximations to the velocities for large t

If an individual has to pass through all the intervening colonies $i\pm1,\ldots,j\mp1$ in order to travel from colony i to colony j, we may regard migration as being between nearest neighbours only. Let us therefore consider the migration distribution

$$h_1 = \nu_1/\nu \quad ; \quad h_{-1} = \nu_2/\nu \quad ; \quad h_i = 0 \quad \text{otherwise} \quad .$$

Then as (2.2) implies

$$\nu = \nu_1 + \nu_2 \ ,$$

expression (2.6) may be written as

$$(3.1) \qquad \sum_{i=-\infty}^{\infty} m_i(t) z^i = e^{(\lambda-\mu-\nu)t} \exp\{t(\nu_1 z + \nu_2 z^{-1})\} \ .$$

Expanding the right hand side of (3.1) in powers of z and comparing coefficients of z^i on both sides of the resulting expression we get

$$(3.2) \qquad m_i(t) = e^{(\lambda-\mu-\nu)t} (\nu_1 t)^i \sum_{k=0}^{\infty} \frac{(t\sqrt{(\nu_1\nu_2)})^{2k}}{(i+k)!k!} \qquad (-\infty < i < \infty) \ .$$

Alternatively, a comparison of result (3.1) with (8.2.9) gives the equivalent representation

$$(3.3) \qquad m_i(t) = e^{(\lambda-\mu-\nu)t} (\nu_1/\nu_2)^{i/2} I_i(2t\sqrt{(\nu_1\nu_2)}) \qquad (-\infty < i < \infty)$$

where $I_i(x)$ denotes the modified Bessel function of the first kind. Note that if individuals may migrate one step to the right only, so that $\nu_2=0$, expression (3.2) simplifies to

$$(3.4) \qquad m_i(t) = e^{(\lambda-\mu-\nu_1)t} (\nu_1 t)^i/(i!) \qquad (i=0,1,2,\ldots) \ .$$

If $\nu_1=\nu_2$ so that there is no 'migration gradient', expression (3.2) reduces to Bailey's result (8.1.3) for the situation in which $a_0=1$ and $a_i=0$ $(i\neq 0)$.

As t increases, the population will continue to spread over an ever-increasing number of colonies if the net growth rate $\lambda-\mu$ is positive. The process may, therefore, be envisaged as a population wave spreading over the integer axis. Let us define the position i^* of the wavefront at time t by means of the condition

that

$$m_{i*}(t) = K \tag{3.5}$$

where K is some assigned constant.

Let us assume that subject to this definition the wavefront possesses an asymptotic velocity c, so that

$$i^* \sim ct \tag{3.6}$$

for $t \geq t_0$ where t_0 is suitably large. Putting $i=ct$ in (3.3) we obtain

$$m_{ct}(t) = e^{(\lambda-\mu-\nu_1-\nu_2)t}(\nu_1/\nu_2)^{ct/2} I_{ct}(2t\sqrt{(\nu_1\nu_2)}) \; . \tag{3.7}$$

But it follows from result (9.7.7) of Abramowitz and Stegun (1965) that for large x

$$I_x(xy) \sim (2\pi x)^{-\frac{1}{2}}(1+y^2)^{-\frac{1}{4}} e^{\eta x} \tag{3.8}$$

where

$$\eta = (1+y^2)^{\frac{1}{2}} + \log\{y/[1+(1+y^2)^{\frac{1}{2}}]\} \; . \tag{3.9}$$

Hence (3.7) may be expressed in the asymptotic form

$$m_{ct}(t) \sim (2\pi ct)^{-\frac{1}{2}}(1+4\nu_1\nu_2/c^2)^{-\frac{1}{4}}(\nu_1/\nu_2)^{ct/2} e^{\eta ct} e^{(\lambda-\mu-\nu_1-\nu_2)t} \tag{3.10}$$

where

$$\eta = (1+4\nu_1\nu_2/c^2)^{\frac{1}{2}} + \log\{2(\nu_1\nu_2)^{\frac{1}{2}}/[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]\} \; . \tag{3.9a}$$

Putting

$$m_{ct}(t) = K$$

and taking logs on both sides of (3.10) we get

$$\text{const.} \sim -\tfrac{1}{2}\log(t) + t[c\eta + (\lambda-\mu-\nu_1-\nu_2) + \tfrac{1}{2}c\log(\nu_1/\nu_2)] \; . \tag{3.11}$$

Thus for large t

$$\nu_1+\nu_2+\mu-\lambda = (c^2+4\nu_1\nu_2)^{\frac{1}{2}} - c\log\{[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]/(2\nu_1)\} \; . \tag{3.12}$$

The required asymptotic velocity c is therefore a root of equation (3.12).

Equation (3.12) was also derived by substituting for $i=ct$ in expression (3.2) and putting

$$Y_k(t) = \text{const.} \quad , \tag{3.13}$$

where $Y_k(t)$ denotes the term of maximum order in t in the summation. This method does not involve using the asymptotic result (3.8), but relies instead on the rather dubious assumption that conditions (3.5) and (3.13) are equivalent.

Suppose $\lambda > \mu$ so that the total mean population size is increasing. Then the LHS of expression (3.12) is equal to

$$\nu_1+\nu_2+\mu-\lambda < \nu_1+\nu_2 \quad .$$

In order to examine the RHS put

$$x(c) = (c^2+4\nu_1\nu_2)^{\frac{1}{2}} - c\log\{[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]/(2\nu_1)\} \quad . \tag{3.14}$$

As $c \to +\infty$

$$x(c) \to c - c\log(c/\nu_1) \to -\infty \quad ,$$

whilst as $c \to -\infty$

$$\begin{aligned} x(c) &\to c - c\log\{(-c/2\nu_1)[-1+(1+4\nu_1\nu_2/c^2)^{\frac{1}{2}}]\} \\ &= c - c\log\{(-c/2\nu_1)[-1+1+2\nu_1\nu_2/c^2 - \ldots]\} \\ &\to c - c\log \nu_2 + c\log(-c) \\ &\to -\infty \quad . \end{aligned}$$

Moreover,

$$x(0) = 2(\nu_1\nu_2)^{\frac{1}{2}} \quad . \tag{3.15}$$

Differentiating (3.14) to find the turning points we get

$$\begin{aligned} dx/dc &= -\log\{(1/2\nu_1)[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]\} \\ d^2x/dc^2 &= -(c^2+4\nu_1\nu_2)^{-\frac{1}{2}} \quad . \end{aligned}$$

Thus $dx/dc = 0$ when

$$c + (c^2+4\nu_1\nu_2)^{\frac{1}{2}} = 2\nu_1 \quad ,$$

i.e. when

$$c = \nu_1 - \nu_2 \quad ,$$

and this value corresponds to a maximum as

$$d^2x/dc^2 = -1/(\nu_1+\nu_2) < 0 \quad \text{at} \quad c = \nu_1-\nu_2 \quad .$$

But

$$x(\nu_1-\nu_2) = (\nu_1+\nu_2) - (\nu_1-\nu_2)\log\{(1/2\nu_1)[(\nu_1-\nu_2) + (\nu_1+\nu_2)]\}$$

$$= \nu_1+\nu_2$$

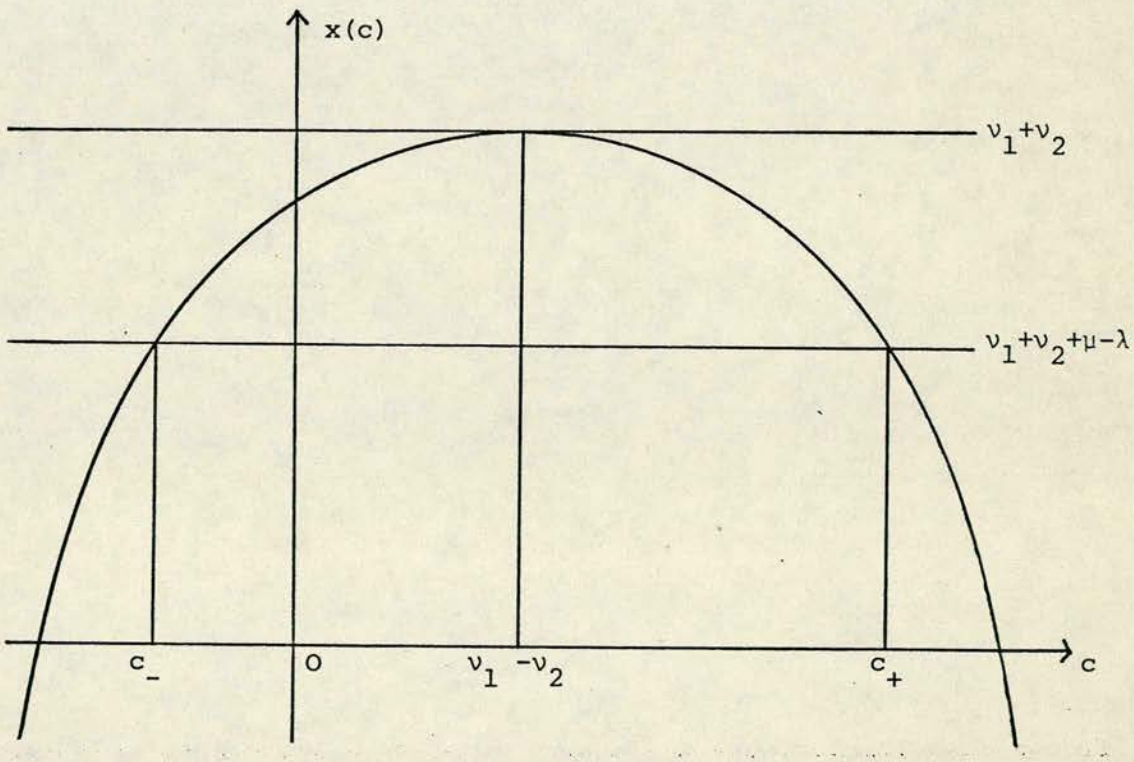
$$> \text{LHS of expression (3.12).}$$

Hence we have the situation as shown in figure 9.1. Equation (3.14) clearly has two real roots c_- and c_+ which correspond to the

Figure 9.1

The wavefront velocities c_+ and c_- determined from the equation $x(c) = \nu_1+\nu_2+\mu-\lambda$

(For explanation see text).



velocities of the left and right wavefronts respectively. Two distinct wavefronts exist because the population spreads along a line, and these wavefronts need not necessarily move in opposite directions.

Suppose we have a drift to the right, so that $\nu_1 > \nu_2$. Then using (3.15), we see from figure 9.1 that $c_- = 0$ when

$$x(0) = 2(\nu_1\nu_2)^{\frac{1}{2}} = \nu_1+\nu_2+\mu-\lambda$$

i.e. when

$$\lambda-\mu = (\nu_1^{\frac{1}{2}} - \nu_2^{\frac{1}{2}})^2 \quad . \tag{3.16}$$

Furthermore, if $\nu_1 > \nu_2$ the left hand boundary moves asymptotically towards $+\infty$ or $-\infty$ according as

$$\lambda-\mu < (\nu_1^{\frac{1}{2}} - \nu_2^{\frac{1}{2}})^2$$

or

$$\lambda-\mu > (\nu_1^{\frac{1}{2}} - \nu_2^{\frac{1}{2}})^2 \quad ,$$

respectively. If $\nu_1 < \nu_2$ expression (3.16) still holds, but the above inequalities, which now refer to the right hand boundary, are reversed.

A possible objection to the above approach is that on substituting from (3.12) into (3.10) we have

$$m_{ct}(t) \sim (2\pi ct)^{-\frac{1}{2}}(1+4\nu_1\nu_2/c^2)^{-\frac{1}{4}} \tag{3.17}$$

which appears to contradict condition (3.5). However, (3.5) is far more stringent than is really necessary. In fact, provided we employ a condition of the form

$$m_{i*}(t) = Kt^{\alpha} \tag{3.18}$$

for some constants $K > 0$ and α, we obtain the same velocities c_+ and c_-. For if we use (3.18) in place of (3.5), equation (3.11) becomes

$$(\alpha/t)\log(t) \sim -(1/2t)\log(t) + [c\eta + (\lambda-\mu-\nu_1-\nu_2) + \tfrac{1}{2}c\log(\nu_1/\nu_2)]$$

i.e. $$c\eta + (\lambda-\mu-\nu_1-\nu_2) + \tfrac{1}{2}c\log(\nu_1/\nu_2) = 0 \ ,$$

which leads to equation (3.12) as before.

9.3.ii One-way migration

If $\nu_2=0$, so that migration may occur only to the right, the above argument becomes far simpler. For the generating function (3.1) yields

(3.4) $$m_i(t) = e^{(\lambda-\mu-\nu_1)t}(\nu_1 t)^i/(i!) \qquad (i=0,1,2,\ldots),$$

which has a Poisson form as opposed to the Bessel form of (3.3). Put

$$m_i(t) = \text{const.} \ ,$$

take logs of both sides and apply Stirling's formula

(3.19) $$i! \sim (2\pi)^{\frac{1}{2}} e^{-i} i^{i+\frac{1}{2}} \ .$$

This procedure yields

$$(\nu_1+\mu-\lambda)t \sim i\log(\nu_1 t/i) + i - \tfrac{1}{2}\log(i) \ .$$

Letting $i=ct$ and considering i and t large as above we have

(3.20) $$\nu_1+\mu-\lambda = c\log(\nu_1/c) + c \ .$$

The required velocity c is therefore a root of equation (3.20). Exactly the same result is obtained if we place $\nu_2=0$ in the more general expression (3.12).

We may now apply a parallel argument to that following expression (3.12). Denote

(3.21) $$x(c) = c\log(\nu_1/c) + c \ ,$$

and suppose $\lambda > \mu$ so that the total mean population size is increasing. Then the LHS of expression (3.20) is equal to

$$\nu_1+\mu-\lambda < \nu_1 \quad .$$

Now as $\nu_2=0$, all colonies sited in positions $i < 0$ are permanently empty, and so c must lie in the range $[0,\infty)$. Clearly

$$x(0) = x(\nu_1 e) = 0$$

with

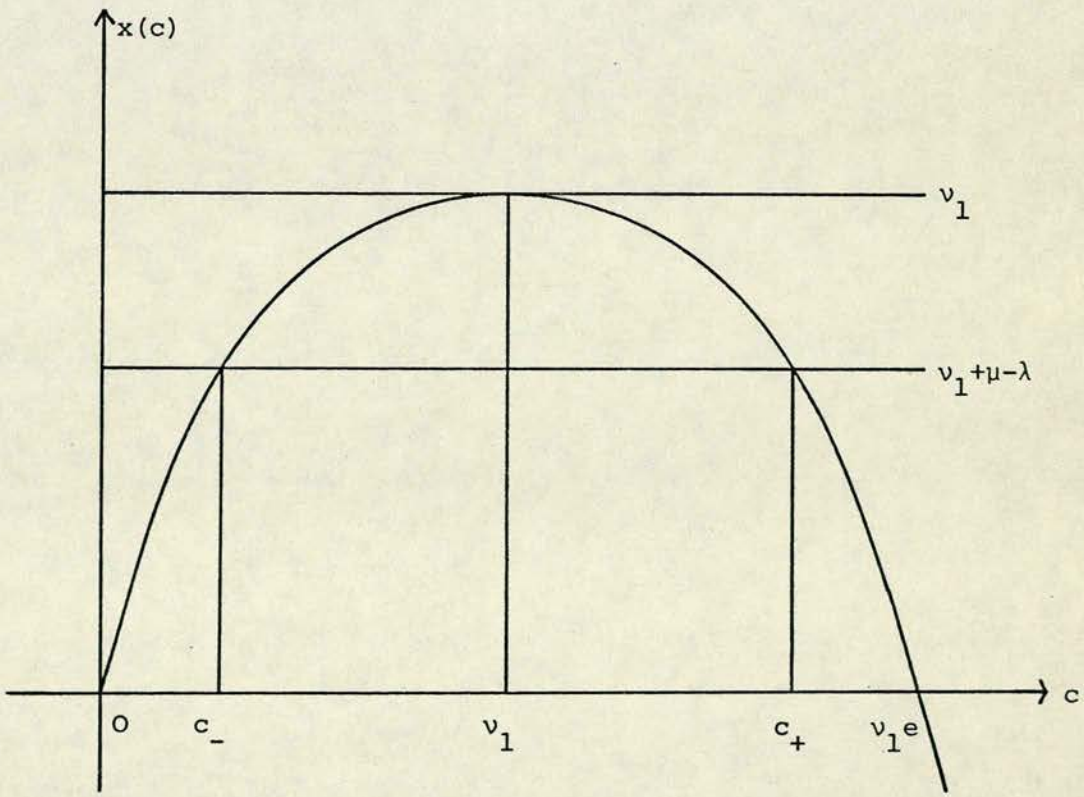
$$x(c) \to -\infty \text{ as } c \to \infty ,$$

and on differentiating (3.21) with respect to c we see that $x(c)$ reaches its maximum value ν_1 when $c = \nu_1$. Hence we have the situation as shown in figure 9.2.

Figure 9.2

The wavefront velocities c_+ and c_- determined from the equation $x(c) = \nu_1+\mu-\lambda$: one-way migration

(For explanation see text).



(i) If $\nu_1+\mu-\lambda>0$, equation (3.20) clearly has two real roots c_- and c_+ , corresponding to the velocities of the left and right wavefronts respectively, where

(3.22) $$0<c_-<\nu_1<c_+<\nu_1 e \quad .$$

(ii) If $\nu_1+\mu-\lambda=0$ then $c_-=0$ and $c_+=\nu_1 e$, whilst if $\nu_1+\mu-\lambda<0$ equation (3.20) has a unique real root $c_+>\nu_1 e$. In both of these cases $m_i(t)\to\infty$ as $t\to\infty$ for all $i=0,1,2,\ldots,$ and so only the right wavefront exists.

9.3.iii Epidemiological examples of velocities

In order that the reader may gain some idea of the range of velocities observed in practice, we present the following three epidemiological examples.

In Glasgow during the early part of the present century annual or biennial outbreaks of measles crossed the city at a rate of about 1 km./month. This rate of advance was so slow that in some winters not all of the city was affected, and this usually led to a more severe outbreak the following year (Halliday (1929), Bailey (1957)). In 1348 and 1349 the Black Death travelled northwards across Western Europe at around 50-100 kms./month (Ziegler (1969)). Stem rust, which is caused by a fungus with light airborne spores, has been known to travel from Texas to Canada at an average rate of over 500 kms./month in favourable spring weather (Stakman and Harrar (1957)).

9.4 Velocities for Nearest-Neighbour Migration based on Tail Aggregates

Expression (3.12) has been determined under the definition that the required velocity c at time t is the relevant root of the

equation

$$m_{ct}(t) = \text{const.} \tag{4.1}$$

An alternative definition of velocity is given by $c=i^*/t$, where i^* is given by

$$\sum_{j=i^*}^{\infty} m_j(t) = \text{const.} \tag{4.2}$$

Here the whole tail of the population spread is taken into account instead of the mean number of individuals at a single colony alone. The question arises as to whether the velocities derived from conditions (4.1) and (4.2) are the same.

Summing expressions (3.2) we have

$$\sum_{n=0}^{\infty} m_{n+i}(t) = e^{(\lambda-\mu-\nu)t} \sum_{n=0}^{\infty} (\nu_1 t)^{n+i} \sum_{k=0}^{\infty} \frac{(t\sqrt{(\nu_1\nu_2)})^{2k}}{(n+i+k)!k!} . \tag{4.3}$$

Each term on the right hand side is non-negative, and on placing $z=1$ in (3.1) and noting that $m_i(t) \geq 0$ for all $-\infty < i < \infty$ and $t \geq 0$ we see that

$$\sum_{n=0}^{\infty} m_{n+i}(t) < \sum_{i=-\infty}^{\infty} m_i(t) = e^{(\lambda-\mu)t} .$$

Thus the summation signs in expression (4.3) may be interchanged, whence

$$\sum_{n=0}^{\infty} m_{n+i}(t) = e^{(\lambda-\mu-\nu)t} (\nu_1 t)^i \sum_{k=0}^{\infty} \frac{(t\sqrt{(\nu_1\nu_2)})^{2k}}{(i+k)!k!} \sum_{n=0}^{\infty} \frac{(\nu_1 t)^n (i+k)!}{(n+i+k)!} . \tag{4.4}$$

Denote

$$Q_k = \sum_{n=0}^{\infty} \frac{(\nu_1 t)^n (i+k)!}{(n+i+k)!}$$

$$= 1 + \sum_{n=1}^{\infty} \frac{(\nu_1 t)^n}{(i+k+1)\ldots(i+k+n)} \tag{4.5}$$

and replace i by $c_+ t$. Then

$$Q_k = 1 + \sum_{n=1}^{\infty} \frac{\nu_1^n}{[c_+ + (k+1)/t] \ldots [c_+ + (k+n)/t]} \quad .$$

Thus under the assumption that $c_+ > \nu_1$, we have that

$$Q_k < \sum_{n=0}^{\infty} (\nu_1/c_+)^n = (1-\nu_1/c_+)^{-1}$$

for all $t \geq 0$ and $k=0,1,2,\ldots$. Substituting this inequality into (4.4) we get

$$(4.6) \qquad m_i(t) < \sum_{n=0}^{\infty} m_{n+i}(t) < m_i(t)(1-\nu_1/c_+)^{-1} \qquad (c_+ > \nu_1) \quad ,$$

so conditions (4.1) and (4.2) are equivalent. A corresponding result may be derived for the left-hand wavefront.

It follows from figure 9.2 that if $\nu_2 = 0$, so that migration occurs to the right only, $c_+ > \nu_1$ provided, of course, that $\lambda > \mu$. However, if $\nu_2 > 0$ we see from figure 9.1 that c_+ may take any value in the range $(\nu_1 - \nu_2, \infty)$. Thus for certain parameter values we have $\nu_1/c_+ > 1$ and the argument above breaks down.

Indeed, $Q_k(t) \to \infty$ as $t \to \infty$ for all $k=0,1,2,\ldots$. For (4.5) gives

$$Q_k > \sum_{n=0}^{\infty} \left(\frac{\nu_1 t}{i+k+n} \right)^n$$

which on putting $i = c_+ t$

$$= \sum_{n=0}^{\infty} \left(\frac{\nu_1}{c_+ + (k+n)/t} \right)^n \quad .$$

Assume $\nu_1 > c_+$ and choose $N(t)$ such that for fixed k

$$c_+ + (k+N)/t \leq \nu_1 < c_+ + (k+N+1)/t \quad .$$

Then

$$Q_k > \sum_{n=0}^{N} \left(\frac{\nu_1}{c_+ + (k+n)/t} \right)^n$$

$$> N+1$$

$$> (\nu_1 - c_+)t - k \quad .$$

However, all is not yet lost. For if we could prove that

$$Q_k = o(t^d) \qquad (k=0,1,2,\ldots) \tag{4.7}$$

for some constant d, an argument similar to that following (3.18) would show that conditions (4.1) and (4.2) are equivalent. Unfortunately, by writing (4.5) in the form

$$Q_k = (i+k)!(\nu_1 t)^{-(i+k)}\left[e^{\nu_1 t} - \sum_{n=0}^{i+k-1} (\nu_1 t)^n/n!\right] ,$$

and noting that if $\nu_1 t > i+k-1$ then

$$\max_{0 \le n \le i+k-1} [(\nu_1 t)^n/n!] = (\nu_1 t)^{i+k-1}/(i+k-1)! ,$$

we may show that for every integer $k \ge 0$ there exists a $T_k > 0$ such that

$$Q_k > e^{bt} \qquad (t > T_k) \tag{4.8}$$

for some constant $b > 0$. However, $T_k \to \infty$ as $k \to \infty$ and so result (4.8) does not lead to a proof that conditions (4.1) and (4.2) yield different velocities but merely shows that this present combinatorial approach is no longer worth pursuing.

In the following section we shall prove that in the vicinity of the wavefront the $m_i(t)$ decrease at least geometrically as i increases and t remains fixed. Thus it seems intuitively reasonable that the two conditions (4.1) and (4.2) should yield identical velocities. Mollison has proved this (at present unpublished) by using an analytic as opposed to a combinatorial approach.

9.5 The Effect of Migration Rate on Velocity

Although it is not possible to obtain an explicit solution to equation (3.12) in closed form, numerical solutions are easily derived by the use of, for example, Newton's rule. This states that

if α is an approximation to a root of $f(z)=0$, then in general a better approximation is $\alpha+\delta\alpha$, where

(5.1) $$\delta\alpha = -f(\alpha)/f'(\alpha) \quad .$$

For simplicity let us restrict our attention to one-way migration ($\nu_1=\nu$; $\nu_2=0$). Then the velocity of the right-hand wavefront is given by the largest positive root of equation (3.20), namely

(5.2) $$f(c) = c\log(\nu/c) + c + \xi = 0$$

where $\xi=\lambda-\mu-\nu$. Differentiating (5.2) we get

$$f'(c) = \log(\nu/c)$$

and so

$$f(c)/f'(c) = c + (c+\xi)/\log(\nu/c) \quad .$$

Thus if $\{c_n\}$ is a sequence of approximations generated by (5.1) which converges to the velocity of propagation c as $n\to\infty$,

(5.3) $$c_{n+1} = (c_n+\xi)/\log(c_n/\nu) \quad .$$

An inspection of figure 9.2 suggests the general starting rule

$$c_0 = \begin{cases} 2\nu : \xi < 0 \\ 3\nu : \xi \doteq 0 \\ 4\nu : \xi > 0 \end{cases} \quad .$$

The above technique was used to generate values of c for the parameter values

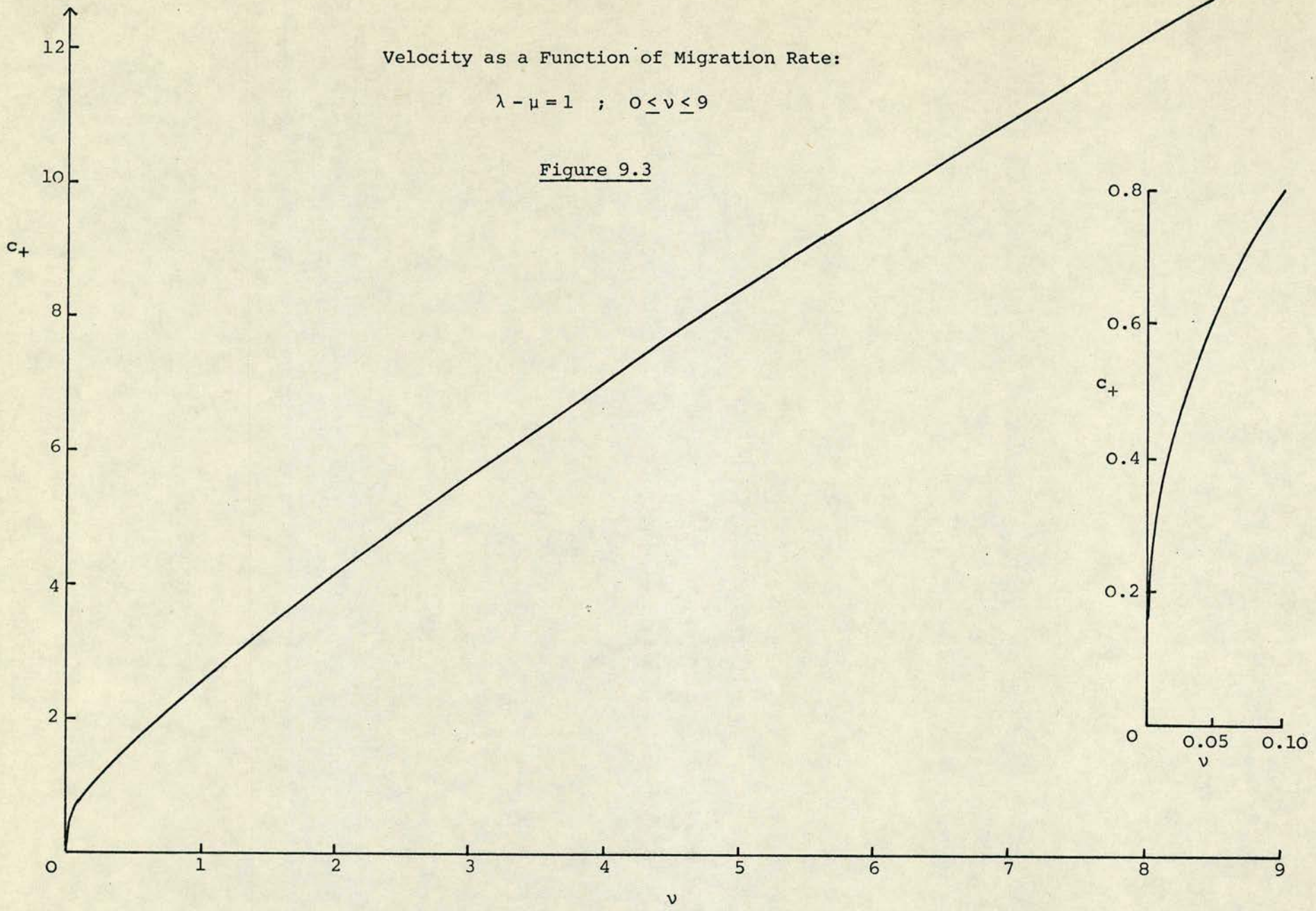
$$\lambda-\mu = 1 \quad \text{and} \quad \nu = 0.01(0.01)0.10 \; ; \quad 0.1(0.1)9.0 \; ,$$

and the curves corresponding to the two ranges of ν are plotted in figure 9.3. From a starting value of $c_0=2\nu$, at most 8 and often as few as 4 iterations were required to achieve an accuracy of within 10^{-5}. The graph of c_+ against ν rises extremely sharply as ν

Velocity as a Function of Migration Rate:

$\lambda - \mu = 1 \quad ; \quad 0 \leq \nu \leq 9$

Figure 9.3



begins to move away from zero, and then as ν becomes larger it quickly flattens out. Now as ν increases, the probability that a given event is a migration tends to 1, and so for large values of ν the effect of the fixed growth rate $\lambda-\mu$ should be negligible. An inspection of figure 9.2 leads us to the conjecture that $c_+/\nu \to 1$ as $\nu \to \infty$.

This conjecture is easily proved. For write

$$c/\nu = a_0 + a_1\nu^d + a_2\nu^{2d} + \dots \tag{5.4}$$

where d and a_i $(i=0,1,2,\dots)$ are unknown constants, and place (5.4) into(5.2), namely

$$-(c/\nu)\log(c/\nu) + (c/\nu) + (\lambda-\mu)/\nu - 1 = 0 \quad . \tag{5.5}$$

Expanding $\log(c/\nu)$ in powers of ν^d and writing down the constant term we obtain

$$-a_0\log(a_0) + a_0 - 1 = 0 \tag{5.6}$$

which has a solution $a_0=1$. But (5.6) corresponds exactly to expression (3.20) if $a_0=c/\nu$ and $\lambda=\mu$, whence it follows from figure 9.2 that $a_0=1$ is the unique solution.

The remaining terms in the expansion of (5.5) reduce to

$$-\tfrac{1}{2}a_1^2\nu^{2d} + (\lambda-\mu)\nu^{-1} + o(\nu^{2d}) = 0 \quad .$$

Thus we may take

$$d = -\tfrac{1}{2} \quad \text{and} \quad a_1 = \{2(\lambda-\mu)\}^{\frac{1}{2}}$$

which gives

$$c/\nu = 1 + \nu^{-\frac{1}{2}}\{2(\lambda-\mu)\}^{\frac{1}{2}} + o(\nu^{-\frac{1}{2}}) \quad . \tag{5.7}$$

The function c/ν clearly approaches the asymptotic value 1 extremely slowly. For example, with $\lambda-\mu=1$,

$$c_+ = 12.321 \quad \text{and} \quad 13.564 \quad \text{when } \nu = 8 \text{ and } 9 \ ,$$

respectively, which gives a gradient of 1.243 between the two points

(see figure 9.3). Even when ν is as large as 200, expression (5.7) yields $c/\nu \sim 1.10$.

9.6 The Wave Shape of First-Order Moments for Nearest-Neighbour Migration

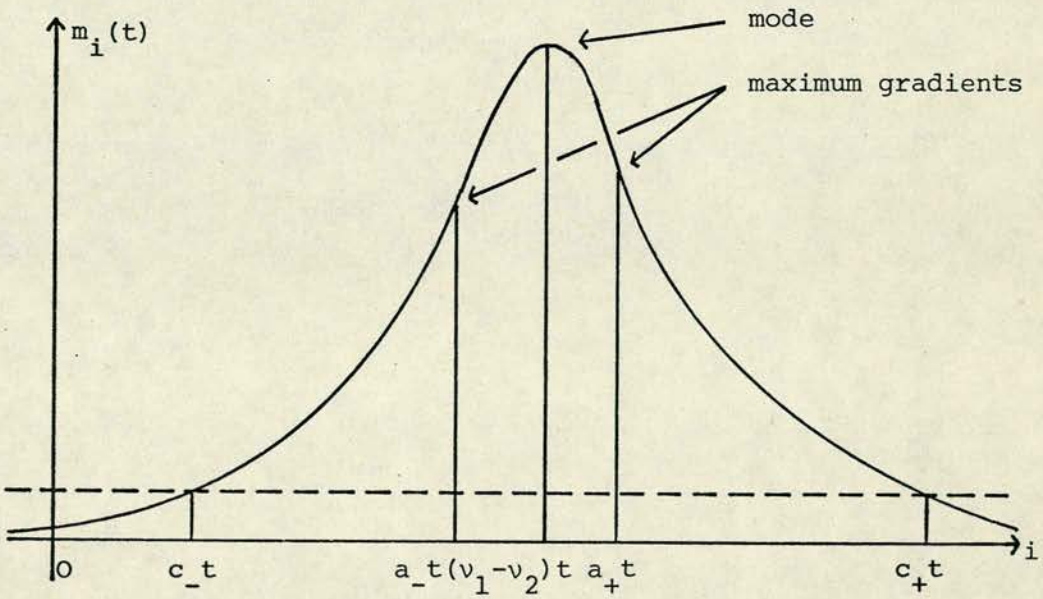
9.6.i First-order moments near the wavefront

Consider a smooth curve drawn through the points $m_i(t)$ where $t \geq 0$ is assumed fixed and $-\infty < i < \infty$. This curve, which is illustrated in figure 9.4, may be thought of as representing a travelling wave. Features of interest include the structure of the wave at the wavefronts themselves and the position and velocity of the mode and points of steepest slope.

In order to investigate the structure of the wavefront we shall consider the function $m_{c_+ t+k}(t)/m_{c_+ t}(t)$. This is the ratio of

Figure 9.4

The shape of the function $m_i(t)$ for fixed t and varying i (For explanation see text).



the mean number of individuals in two colonies, situated near the wavefront, which are a distance k apart. For typographical convenience let us write c for c_+. Putting $i=ct+k$ in (3.3) we have

$$m_{ct+k}(t) = e^{(\lambda-\mu-\nu)t}(\nu_1/\nu_2)^{\frac{1}{2}(ct+k)} I_{ct+k}[(ct+k)\{2(\nu_1\nu_2)^{\frac{1}{2}}(c+k/t)^{-1}\}]$$

which from result (3.8) may be approximated for large t by

$$(6.1)\quad m_{ct+k}(t) \sim e^{(\lambda-\mu-\nu)t}(\nu_1/\nu_2)^{\frac{1}{2}(ct+k)} e^{(ct+k)\eta'} [2\pi(ct+k)]^{-\frac{1}{2}} \times [1 + 4\nu_1\nu_2(c+k/t)^{-2}]^{-\frac{1}{4}}$$

where

$$(6.2)\quad \eta' = [1+4\nu_1\nu_2(c+k/t)^{-2}]^{\frac{1}{2}} - \log\{[(c+k/t) + [(c+k/t)^2 + 4\nu_1\nu_2]^{\frac{1}{2}}] \div 2(\nu_1\nu_2)^{\frac{1}{2}}\} .$$

Dividing expression (6.1) by (3.10) we get

$$(6.3)\quad m_{ct+k}(t)/m_{ct}(t) \sim (\nu_1/\nu_2)^{\frac{1}{2}k} e^{(ct+k)\eta'-ct\eta}$$

where η is given by (3.9a) and η' by (6.2). After a little algebra the exponential term reduces to the asymptotic value

$$\{[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]/2(\nu_1\nu_2)^{\frac{1}{2}}\}^{-k}$$

whence

$$(6.4)\quad m_{ct+k}(t)/m_{ct}(t) \sim \{2\nu_1/[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]\}^k .$$

Now we see from figure 9.1 that $c_+ > \nu_1-\nu_2$. Hence as $c_+ + (c_+^2+4\nu_1\nu_2)^{\frac{1}{2}}$ is an increasing function of c_+,

$$c_+ + (c_+^2 + 4\nu_1\nu_2)^{\frac{1}{2}} > (\nu_1-\nu_2) + (\nu_1 + \nu_2) = 2\nu_1 .$$

Therefore, in the vicinity of the wavefront, the $m_i(t)$ asymptotically decrease geometrically at rate

$$2\nu_1/[c_+ + (c_+^2+4\nu_1\nu_2)^{\frac{1}{2}}]$$

as i increases.

9.6.ii The modal first-order moments

To determine the value of i which maximises the function $m_i(t)$, where t is considered fixed, we may either treat the integer i as a continuous variable and differentiate $m_i(t)$ with respect to it, or we may find that value of i for which

$$m_{i-1}(t) < m_i(t) > m_{i+1}(t)$$

(care must be taken to ensure that this discrete search yields a global and not a local maximum). Both of these approaches produce unmanageable expressions and so we shall employ a less rigorous technique.

Denote

$$Y_k(t) = \frac{[t\sqrt{(\nu_1\nu_2)}]^{2k}}{(i+k)!k!}$$

to be the k^{th} term of the summation on the right hand side of expression (3.2). Then on using Stirling's formula (3.19) we have

(6.5) $$\log[Y_k(t)] \sim i + 2k + 2k\log[t\sqrt{(\nu_1\nu_2)}] - (i+k+\tfrac{1}{2})\log(i+k) - (k+\tfrac{1}{2})\log(k) \quad .$$

For the purpose of differentiation consider the integer k to be replaced by a continuous variable. Then differentiation of both sides of (6.5) with respect to k yields (for k large)

$$\frac{\partial}{\partial k}[\log[Y_k(t)]] \sim 2\log[t\sqrt{(\nu_1\nu_2)}] - \log[k(i+k)]$$

which equals zero at

(6.6) $$\nu_1\nu_2 t^2 = k(i+k) \quad .$$

[A comparable result is obtained if we seek the value of k which satisfies the inequalities

$$\frac{(t\sqrt{(\nu_1\nu_2)})^{2(k-1)}}{(i+k-1)!(k-1)!} < \frac{(t\sqrt{(\nu_1\nu_2)})^{2k}}{(i+k)!k!} > \frac{(t\sqrt{(\nu_1\nu_2)})^{2(k+1)}}{(i+k+1)!(k+1)!} \quad] \ .$$

Equation (6.6) has roots

$$k_1, k_2 = \tfrac{1}{2}[-i \pm (i^2+4\nu_1\nu_2 t^2)^{\frac{1}{2}}] \ .$$

But $k \geq 0$, and so for fixed $t \geq 0$ the maximum value of $Y_k(t)$, denoted by $Y_{max}(t)$, occurs when

(6.7) $$k = k_1 = \tfrac{1}{2}[-i+(i^2+4\nu_1\nu_2 t^2)^{\frac{1}{2}}] \ .$$

Let us therefore write expression (3.2) in the form

(6.8) $$m_i(t) = e^{(\lambda-\mu-\nu)t}(\nu_1 t)^i Y_{max}(t)[1+o(1)] \ .$$

Define the mode to be at position i at time $t \geq 0$ if

(6.9) $$m_{i-\frac{1}{2}}(t) = m_{i+\frac{1}{2}}(t)$$

where i is now taken to be continuous. [Alternatively, consider the inequalities $m_{i-1}(t) < m_i(t) > m_{i+1}(t)$] . Then substituting from (6.8) into (6.9) we get

$$\frac{(\nu_1 t)^{i-\frac{1}{2}}(t\sqrt{(\nu_1\nu_2)})^{2k_1}}{\Gamma(i+k_1+\frac{1}{2})k_1!} = \frac{(\nu_1 t)^{i+\frac{1}{2}}(t\sqrt{(\nu_1\nu_2)})^{2k_1}[1+o(1)]}{\Gamma(i+k_1+3/2)k_1!}$$

i.e.

(6.10) $$i+k_1+\tfrac{1}{2} = \nu_1 t[1+o(1)] \ .$$

Here we denote the gamma-function by

(6.11) $$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt \ .$$

On combining (6.7) and (6.10) we deduce that for large t

$$(i^2/t^2 + 4\nu_1\nu_2) \sim (2\nu_1 - i/t)^2 \ ,$$

and this gives

(6.12) $$i/t = \nu_1 - \nu_2 \ .$$

Thus the asymptotic velocity of the mode is simply the difference between the two migration rates ν_1 and ν_2. This certainly seems intuitively reasonable.

In order to obtain result (6.12) we have replaced the true value of $m_i(t)$ at time $t \geq 0$ by its term of maximum magnitude which is given by (6.8). A partial justification for this approach is provided by the comment immediately after expression (3.12); namely, that the same velocities c_+ and c_- are obtained if the condition $m_i(t) = \text{constant}$ is replaced by $Y_{max}(t) = \text{constant}$.

9.6.iii Points of steepest slope in the wave for one-way migration

It follows from figure 9.4 that the wave possesses two points of steepest slope. We shall derive the velocities of these two points under the assumption of one-way migration ($\nu_2=0$). The gain in information by considering $\nu_2 > 0$ would be far outweighed by an excessive increase in the amount of algebra.

For fixed $t \geq 0$ the modulus of the gradient, $|m_{j-1}(t)-m_j(t)|$, is a maximum when

$$m_{j-1}(t) - m_j(t) \doteq m_j(t) - m_{j+1}(t) \quad .$$

(This condition for determining the 'points of inflexion' is the discrete equivalent of equating the second-order derivative of a continuous function to zero). Thus on substituting for $m_i(t)$ from (3.4) we get

$$\frac{(\nu_1 t)^{j-1}}{(j-1)!} - \frac{2(\nu_1 t)^j}{j!} + \frac{(\nu_1 t)^{j+1}}{(j+1)!} \doteq 0$$

whence

$$j \doteq \tfrac{1}{2}[(2\nu_1 t-1) \pm (1+4\nu_1 t)^{\frac{1}{2}}] \quad .$$

Hence the points of steepest slope are at a_+t and a_-t, where

(6.13) $$a_+ \sim \nu_1 + (\nu_1/t)^{\frac{1}{2}} \quad \text{and} \quad a_- \sim \nu_1 - (\nu_1/t)^{\frac{1}{2}}$$

correspond to their velocities.

9.6.iv The time maxima of first-order moments

On retaining the assumption that $\nu_2=0$ we may deduce from (3.4) that if $\lambda-\mu-\nu_1 \leq 0$ each $m_i(t)$ $(i=0,1,2,\ldots)$ increases without bound as $t\to\infty$. However, if $\lambda-\mu-\nu_1 < 0$ each $m_i(t)$ reaches a maximum value $\tilde{m}_i$ at some time $t=t_i>0$. For on differentiating expression (3.4) we get

$$dm_i(t)/dt = e^{(\lambda-\mu-\nu_1)t}\,\nu_1^i t^{i-1}[(\lambda-\mu-\nu_1)t+i]/(i!) \quad ,$$

and if $\lambda-\mu-\nu_1 < 0$ the right hand side equals zero at $t=t_i$ where

(6.14) $$t_i = i/(\mu+\nu_1-\lambda) \quad .$$

This turning point t_i corresponds to the maximum value of $m_i(t)$, namely

(6.15) $$\tilde{m}_i = e^{-i}(i!)^{-1}[\nu_1 i/(\mu+\nu_1-\lambda)]^i \quad .$$

If we apply Stirling's formula (3.19), expression (6.15) simplifies to

(6.16) $$\tilde{m}_i \sim (2\pi i)^{-\frac{1}{2}}[\nu_1/(\mu+\nu_1-\lambda)]^i \quad .$$

Now $\mu+\nu_1-\lambda > 0$, and so as i tends to infinity the maxima $\tilde{m}_i$ either increase or decrease 'almost geometrically' at rate $\nu_1/(\mu+\nu_1-\lambda)$ according as $\mu+\nu_1 > \lambda > \mu$ or $\lambda < \mu$, respectively. If $\lambda=\mu$ we have

$$\tilde{m}_i \sim (2\pi i)^{-\frac{1}{2}} \quad .$$

From (6.14)

$$i/t_i = (\mu+\nu_1-\lambda) \qquad (\lambda-\mu-\nu_1 < 0) \quad .$$

Thus the maxima $\tilde{m}_i$ may be regarded as generating a wave emanating from the origin which has an asymptotic velocity equal to $(\mu+\nu_1-\lambda)$.

9.7 Birth-Death-Migration Models in Continuous State Space

9.7.i Velocity of propagation in two-dimensional space

Consider a simple continuous state space birth-death-migration process in two dimensions such that an individual in position $\underline{v}$ has birth rate $\lambda(\underline{v})$, death rate $\mu(\underline{v})$ and migration rate $\nu(\underline{u}-\underline{v})$ from $\underline{v}$ to $\underline{u}$. Note that we are assuming that the migration distribution is spatially homogeneous, that is $\nu(\underline{u}-\underline{v})$ depends solely on the vector difference $\underline{u}-\underline{v}$ and not on the absolute values of $\underline{u}$ and $\underline{v}$ themselves. Let $p_t(\underline{v})$ denote the probability of extinction of the process by time t given that the initial population at time $t=0$ consists of a single individual at $\underline{v}$. Then consideration of all possible transitions in the first infinitesimal time interval $(0,\delta t)$ yields

$$(7.1)\quad dp_t(\underline{v})/dt = \lambda(\underline{v})p_t^2(\underline{v}) - p_t(\underline{v})\left[\lambda(\underline{v})+\mu(\underline{v})+\int\nu(\underline{u}-\underline{v})d\underline{u}\right] + \mu(\underline{v}) + \int\nu(\underline{v}-\underline{u})p_t(\underline{u})\cdot d\underline{u} \quad .$$

If λ and μ are independent of the position vector $\underline{v}$ we may assume that the probability of extinction $p_t(\underline{v})$ is independent of $\underline{v}$ for all $\underline{v}$, and so equation (7.1) reduces to

$$dp_t(\underline{v})/dt = \lambda p_t^2(\underline{v}) - (\lambda+\mu)p_t(\underline{v}) + \mu$$

which is the well-known equation for the probability of ultimate extinction of a simple one-colony birth and death process. Thus

$$(7.2)\qquad p_t(\underline{v}) \equiv p(t) = (\mu-\mu e^{(\lambda-\mu)t})(\mu-\lambda e^{(\lambda-\mu)t})^{-1}$$

(Cox and Miller (1965), p.166).

Denote the average number of individuals at time t in the element of area $d\underline{r}$ by $f(\underline{r};t)d\underline{r}$. Then, considering all possible changes in the infinitesimal time interval $(t,t+\delta t)$, we have

(7.3) $$f(\underline{r};t+\delta t) = f(\underline{r};t) + f(\underline{r};t)\,[\lambda(\underline{r})-\mu(\underline{r})-\nu]\,\delta t + \delta t \int \nu(\underline{r}-\underline{s})\,f(\underline{s})\,d\underline{s}$$

where

(7.4) $$\nu = \int \nu(\underline{u})\,d\underline{u}$$

denotes the overall rate of migration. All integrals are taken over the whole space $\mathfrak{R}_2$. Letting $\delta t \to 0$ in (7.3) we obtain the equation

(7.5) $$\partial f(\underline{r};t)/\partial t = [\lambda(\underline{r})-\mu(\underline{r})-\nu]\,f(\underline{r};t) + \int \nu(\underline{r}-\underline{s})\,f(\underline{s};t)\,d\underline{s}\ .$$

Denote the Fourier transforms of $f(\underline{r};t)$ and $\nu(\underline{r})$ by

$$f^*(\underline{\theta};t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(\theta_1 x+\theta_2 y)}\, f(x,y;t)\,dxdy$$

and

$$\nu^*(\underline{\theta}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(\theta_1 x+\theta_2 y)}\, \nu(x,y)\,dxdy$$

where $\underline{\theta} = (\theta_1,\theta_2)$. Then taking the Fourier transformation of both sides of equation (7.5) and integrating we obtain, for $\lambda(\underline{r})=\lambda$ and $\mu(\underline{r})=\mu$,

(7.6) $$f^*(\underline{\theta};t) = \text{const.}\exp\{(\lambda-\mu-\nu)t\}\exp\{\nu^*(\underline{\theta})t\}\ .$$

If we assume that at $t=0$ the initial population consists of a single individual at $\underline{r}=\underline{0}$, we have

$$f(x,y;0) = \delta(x)\,\delta(y)\ ,$$

and so

(7.7) $$f^*(\underline{\theta};0) = 1\ .$$

Here $\delta(x)$ denotes the Dirac delta function which has the property that

(7.8) $$\delta(x) = 0 \quad (x\neq 0)\ ;\quad \int_{-\infty}^{\infty}\delta(x)\,dx = 1\ .$$

Combining (7.6) and (7.7) we get

(7.9) $$f^*(\underline{\theta};t) = \exp\{(\lambda-\mu-\nu)t\}\exp\{\nu^*(\underline{\theta})t\} \quad .$$

Suppose that the migration distribution $\nu(\underline{r})$ is bivariate normal with

(7.10) $$\nu(\underline{r}) = (\nu/2\pi\sigma^2)\exp\{-\tfrac{1}{2}(x^2+y^2)/\sigma^2\} \quad .$$

Then for $\underline{\theta} = (\theta_1, \theta_2)$

$$\nu^*(\underline{\theta}) = \nu\exp\{-\tfrac{1}{2}\sigma^2(\theta_1^2+\theta_2^2)\} \quad ,$$

and for small σ^2 expression (7.9) becomes

$$f^*(\underline{\theta};t) = \exp\{(\lambda-\mu-\nu)t\}\exp\{\nu t[1-\tfrac{1}{2}\sigma^2(\theta_1^2+\theta_2^2)+o(\sigma^2)]\} \quad .$$

Hence

$$f^*(\underline{\theta};t) \doteq \exp\{(\lambda-\mu)t\}\exp\{-\tfrac{1}{2}\nu t\sigma^2(\theta_1^2+\theta_2^2)\}$$

which inverts to give

(7.11) $$f(x,y;t) \doteq (2\pi\nu t\sigma^2)^{-1}\exp\{(\lambda-\mu)t\}\exp\{-\tfrac{1}{2}(x^2+y^2)/(\nu t\sigma^2)\} \quad .$$

It is worth noting that if (7.10) holds then expression (7.9) may be written in the form

$$f^*(\underline{\theta};t) = e^{(\lambda-\mu-\nu)t}\sum_{j=0}^{\infty}\frac{(\nu t)^j}{j!}\exp[-\tfrac{1}{2}\sigma^2 j(\theta_1^2+\theta_2^2)]$$

which inverts to give

(7.12) $$f(x,y;t) = e^{(\lambda-\mu-\nu)t}[\delta(x)\,\delta(y) + \sum_{j=1}^{\infty}\frac{(\nu t)^j}{j!}(2\pi j\sigma^2)^{-\frac{1}{2}}\exp\{\frac{-(x^2+y^2)}{2j\sigma^2}\}].$$

Now while $f(\underline{r};t)$ is an average density which includes the zero contribution from cases where total extinction has occurred, the average density $f'(\underline{r};t)$ conditional on no extinction is given by

(7.13) $$f' = f/[1-p(t)] \quad .$$

Let us define the advance of the population wave by the condition that the integral of f' beyond a circle with centre $\underline{0}$ and radius R is to be an assigned constant. Then a comparison of

(7.2), (7.11) and (7.13) shows that for $\lambda > \mu$

$$f'(x,y;t) \sim [\lambda/(\lambda-\mu)](2\pi\nu t\sigma^2)^{-1}\exp\{(\lambda-\mu)t\}\exp\{-\tfrac{1}{2}(x^2+y^2)/(\nu t\sigma^2)\}$$

which, with $z^2=x^2+y^2$, gives

$$\int_R^\infty 2\pi z f'\,dz \sim \lambda[(\lambda-\mu)\nu t\sigma^2]^{-1} e^{(\lambda-\mu)t} \int_R^\infty e^{-z^2/2\nu t\sigma^2} z\,dz$$

$$= [\lambda/(\lambda-\mu)]e^{(\lambda-\mu)t}e^{-R^2/2\nu t\sigma^2} \quad .$$

Hence we require

$$-R^2/(2\nu t\sigma^2) + (\lambda-\mu)t = \text{const.}$$

Put

$$R = vt \ .$$

Then for large R and t we require

$$-v^2/(2\nu\sigma^2) + (\lambda-\mu) = 0 \ ,$$

and so the ultimate velocity of propagation of the wave is given by

(7.14) $$v \sim \sigma[2\nu(\lambda-\mu)]^{\frac{1}{2}} \quad (\lambda > \mu) \ .$$

Skellam (1951) examined the rate of spread of muskrats which were introduced into central Europe in 1905. On five occasions in the succeeding 23 years the area occupied by the expanding population was mapped, so enabling contours to be drawn to show the population's extent at five different times. Although the contours are not circular, there is no evidence of drift, and it seems reasonable to treat the area within each contour as an estimate of $\pi R^2(t)$. Now as the muskrats spread into areas not occupied by potential competitors, it is reasonable to postulate that the muskrat population was increasing exponentially in numbers, whence it follows from (7.14) that the square root of the area within each contour should be linearly related to time. Skellam found that this relationship did, in fact, exist; in a plot of $(\text{area})^{\frac{1}{2}}$ versus

time the five points lie very close to a straight line passing through the origin.

Bartlett (1960) derives a result similar to (7.14) for the spread of an epidemic wave. Here the transition probability rate of infection of a susceptible in the infinitesimal element of area $d\underline{s}$ by an infective in $d\underline{r}$ is assumed to be a function $\lambda(\underline{s}-\underline{r})$ of the vector displacement $\underline{s}-\underline{r}$. In particular he assumes that

$$\lambda(x,y) = (\lambda/2\pi\sigma^2)\exp\{-\tfrac{1}{2}(x^2+y^2)/\sigma^2\} \quad ,$$

where σ is small, so that only local infection occurs. He neglects the movement of susceptibles and replaces them by an average 'density' of n susceptibles per unit area. Under these conditions Bartlett shows that the ultimate velocity of propagation of the epidemic wave is given by

$$\sigma[2n\lambda(n\lambda-\mu)]^{\frac{1}{2}} \qquad (n\lambda > \mu) \quad , \tag{7.15}$$

where μ denotes the death rate of the infectives. Thus, on comparing (7.14) and (7.15), we see that as far as the velocities of the two waves are concerned the epidemic process may be considered as being equivalent to a birth-death-migration process with migration rate $n\lambda$, birth rate $n\lambda$ and death rate μ.

If $\lambda < \mu$, then no such limiting velocity v exists. For

$$f'(x,y;t) \sim (2\pi\nu t\sigma^2)^{-1}(\mu/(\mu-\lambda))e^{-(x^2+y^2)/(2\nu t\sigma^2)}$$

which, with $z^2=x^2+y^2$, gives

$$\int_R^\infty 2\pi z f' dz \sim (\mu/(\mu-\lambda))e^{-R^2/(2\nu t\sigma^2)} \quad .$$

Hence we require

$$R^2/2\nu t\sigma^2 = \text{const.}$$

which, with $R=vt$, gives

$$v = O(t^{-\frac{1}{2}}) \quad . \tag{7.16}$$

Note that $4\pi R^2/t = \text{constant}$, and so the area occupied by the population has a constant rate of growth.

9.7.ii Velocity of propagation in one-dimensional space

We may consider the analagous process in one-dimensional space in a similar manner. Suppose that the migration distribution has a normal form with variance σ^2 and drift m . Then

$$(7.17) \qquad \nu(x) = \nu(2\pi\sigma^2)^{-\frac{1}{2}}\exp\{-(x-m)^2/2\sigma^2\}$$

with Fourier transform

$$(7.18) \qquad \nu^*(\theta) = \nu\exp\{im\theta-\tfrac{1}{2}\sigma^2\theta^2\} \quad .$$

Substituting for $\nu^*(\theta)$ from (7.18) into (7.9) we obtain

$$(7.19) \qquad f^*(\theta;t) = e^{(\lambda-\mu-\nu)t}\exp\{\nu te^{(im\theta-\frac{1}{2}\sigma^2\theta^2)}\} \quad ,$$

and for small m and σ^2 expression (7.19) becomes

$$f^*(\theta;t) \doteq e^{(\lambda-\mu-\nu)t}\exp\{\nu t[1+im\theta-\tfrac{1}{2}\sigma^2\theta^2]\}$$

which inverts to give

$$(7.20) \qquad f(x;t) \doteq (2\pi\nu t\sigma^2)^{-\frac{1}{2}}e^{(\lambda-\mu)t}\exp\{-(x-m\nu t)^2/(2\nu t\sigma^2)\} \quad .$$

The density is therefore a maximum at $x=m\nu t$ when it has value

$$(2\pi\nu t\sigma^2)^{-\frac{1}{2}}\exp\{(\lambda-\mu)t\} \quad .$$

Note that expression (7.19) may be written in the form

$$f^*(\theta;t) = e^{(\lambda-\mu-\nu)t}\sum_{j=0}^{\infty}\frac{(\nu t)^j}{j!}\exp\{j(im\theta-\tfrac{1}{2}\sigma^2\theta^2)\}$$

which inverts to give

$$(7.21) \qquad f(x;t) = e^{(\lambda-\mu-\nu)t}[\delta(x) + \sum_{j=1}^{\infty}\frac{(\nu t)^j}{j!}(2\pi j\sigma^2)^{-\frac{1}{2}}\exp\{\frac{-(x-jm)^2}{2j\sigma^2}\}] \quad .$$

Expressions (7.12) and (7.21) are both mixtures of normal distributions with a Poisson mixing distribution.

Let us define the advance of the population wave by means of

the condition that the density

(7.22) $$f'(R;t) = \text{constant}.$$

Note that identical velocities are obtained if we use the alternative condition that

$$\int_R^\infty f'(x;t)\,dx = \text{constant}.$$

Then on combining expressions (7.2), (7.13), (7.20) and (7.22) we see that for $\lambda > \mu$

$$\text{const.} = t^{-\frac{1}{2}} e^{(\lambda-\mu)t} \exp\{-(R-m\nu t)^2/(2\nu t\sigma^2)\}\ ,$$

which, if we take logs of both sides, becomes

$$\text{const.} = -\tfrac{1}{2}\log(t) + (\lambda-\mu)\,t - (R-m\nu t)^2/(2\nu t\sigma^2)\ .$$

Put

$$R = vt\ .$$

Then for large R and t

$$0 = (\lambda-\mu) - (v-m\nu)^2/(2\nu\sigma^2)\ ,$$

and so the ultimate velocities of the left and right wavefronts, denoted by v_- and v_+ respectively, are given by

(7.23) $$v_+, v_- = m\nu \pm [2\nu\sigma^2(\lambda-\mu)]^{\frac{1}{2}}\ .$$

It follows that the distance between the two wavefronts ultimately increases at rate

$$v_+ - v_- = 2\,[2\nu\sigma^2(\lambda-\mu)]^{\frac{1}{2}}\ ,$$

and if $m > 0$ the left wavefront will move left or right according as $m\nu$ is less than or greater than $[2\nu\sigma^2(\lambda-\mu)]^{\frac{1}{2}}$.

For $\lambda < \mu$ a similar argument shows that

(7.24) $$v_+ = m\nu + O(t^{-\frac{1}{2}}) \quad \text{and} \quad v_- = m\nu + O(t^{-\frac{1}{2}})\ .$$

Thus after a suitable length of time, and conditional on no extinction before or at that time, the distance between the two

wavefronts remains constant.

9.7.iii Comparison of velocities for one-dimensional continuous and nearest-neighbour stepping-stone models

In deriving the velocities c_+ and c_- for the stepping-stone model we assume that a migrating individual may make only discrete jumps of size +1 and -1. However, in evaluating the velocities v_+ and v_- we allow jumps of any size x $(-\infty < x < \infty)$ where the random variable x has the normal form (7.17). We may ask whether the velocities c_+, c_- and v_+, v_- are at all comparable, and if so for what values of the parameters λ, μ, ν_1 and ν_2.

Suppose we have nearest-neighbour migration at rates ν_1 and ν_2 to the right and left respectively. Then in any infinitesimal time period $(t,t+\delta t)$

$$\Pr[\text{an individual moves one step to the right} \mid \text{it migrates}] = \nu_1/(\nu_1+\nu_2)$$

(7.25)

$$\Pr[\text{an individual moves one step to the left} \mid \text{it migrates}] = \nu_2/(\nu_1+\nu_2) \quad ,$$

whilst we see from (7.17) that the corresponding normal distribution is given by

(7.26) $$(2\pi\sigma^2)^{-\frac{1}{2}}\exp\{-(x-m)^2/2\sigma^2\} \quad .$$

Let us equate $m\delta t$ and $\sigma^2\delta t$ to the mean and variance of the distance travelled by an individual, which migrates according to the distribution (7.25), in the infinitesimal time interval $(t,t+\delta t)$.

On writing $\nu=\nu_1+\nu_2$ we see that this procedure yields

$$m\delta t = (\nu_1/\nu)\delta t(1) + (\nu_2/\nu)\delta t(-1) \quad .$$

$$\sigma^2\delta t = (\nu_1/\nu)\delta t(1)^2 + (\nu_2/\nu)\delta t(-1)^2 - (m\delta t)^2$$

which, as $\delta t \to 0$, gives

$$m = (\nu_1-\nu_2)/\nu \quad \text{and} \quad \sigma^2 = 1 .$$

Substituting these values into (7.23) we obtain the normal approximation to the velocities c_+ and c_-, namely

(7.27) $$c_{\text{normal}} = (\nu_1-\nu_2) \pm [2\nu(\lambda-\mu)]^{\frac{1}{2}} .$$

Now we have previously shown that the ultimate velocities c_+ and c_- of the two wavefronts are the roots of the equation

(3.12) $$\nu_1+\nu_2+\mu-\lambda = (c^2+4\nu_1\nu_2)^{\frac{1}{2}} - c\log\{[c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]/(2\nu_1)\} .$$

On putting

(7.28) $$\lambda - \mu = \varepsilon\nu \quad \text{and} \quad \delta = 1 - [c+(c^2+4\nu_1\nu_2)^{\frac{1}{2}}]/(2\nu_1)$$

we may write (3.12) as

(7.29) $$\nu(1-\varepsilon) = (c^2+4\nu_1\nu_2)^{\frac{1}{2}} - c\log(1-\delta)$$

(note that $1-\delta > 0$) . Inverting the expression for δ in (7.28) we get

(7.30) $$c = \nu_1(1-\delta) - \nu_2(1-\delta)^{-1}$$

whence

(7.31) $$(c^2+4\nu_1\nu_2)^{\frac{1}{2}} = \nu_1(1-\delta) + \nu_2(1-\delta)^{-1} .$$

Combining (7.29)-(7.31) we have

(7.32) $$\nu(1-\varepsilon)(1-\delta) = [\nu_1(1-\delta)^2+\nu_2] - [\nu_1(1-\delta)^2-\nu_2]\log(1-\delta) .$$

If we consider ε and δ to be 'small' and disregard terms of greater power than ε and δ^2, expression (7.32) reduces to

(7.33) $$\delta = \pm(2\varepsilon)^{\frac{1}{2}}$$

when $\log(1-\delta)$ is expanded in powers of δ . Now from (7.30)

$$c = \nu_1(1-\delta) - \nu_2(1-\delta)^{-1}$$

$$\doteq \nu_1(1-\delta) - \nu_2(1+\delta) \qquad \text{(small } \delta\text{)}$$

$$= \nu_1 - \nu_2 - \nu\delta \quad .$$

But (7.28) and (7.33) combine to give

$$\delta = \pm[2(\lambda-\mu)/\nu]^{\frac{1}{2}} \quad .$$

Hence

$$c \doteq \nu_1 - \nu_2 \pm [2\nu(\lambda-\mu)]^{\frac{1}{2}} \tag{7.34}$$

which agrees with (7.27).

Thus provided ε and δ are sufficiently small the velocities derived from the stepping-stone model and the continuous state space model are in close agreement. However, it follows from (7.28) and (7.33) that small ε and δ implies

$$0 < \lambda - \mu << \nu \quad .$$

Hence, unless the net growth rate of the process is small in comparison with the sum of the migration rates, 'similar' migration patterns give rise to different velocities in the discrete and continuous cases. That some such difference exists is not surprising, for the distributions (7.25) and (7.26) possess completely different characteristics.

9.8 Stochastic Simulation of a One-Way Nearest-Neighbour Stepping-Stone Model

Although the stepping-stone model described in Chapter 8 is a stochastic process, the associated velocities and waveforms developed in this present chapter have been obtained by considering only the pattern of change in expected colony sizes. In order to assess the relevance of these deterministic results to their stochastic counterparts, a small simulation study was made of the

one-way migration model without death. Here μ and ν_2 were both placed equal to zero.

Except in one respect, the computer program written to simulate this model is directly analagous to the one described in Chapter 7, and so I shall not dwell too long on the computational aspects of it here. The complete program is reproduced in Appendix B; the notation used in it is self-explanatory. The difference between the two programs is that whereas before we had to keep a tally of the numbers in two colonies, this time there are an infinite number of colonies to be considered.

We define an array $N(1),\ldots,N(1000)$, where $N(I)$ denotes the population size of colony $(I-25)$ at a particular instant, and we use the initial conditions

$$N(25) = 1 \quad ; \quad N(I) = 0 \qquad (I \neq 25) \quad .$$

Let R_1, R_2 and R_3 denote different pseudo-random numbers generated by (7.1.3).

(i). Time to next event : lines 28-33

This random variable is exponentially distributed with parameter $Q \equiv (\lambda+\nu_1)(N(25)+\ldots+N(POS))$. Here $POS \geq 25$ denotes the position of the right-hand wavefront, so that $N(POS) > 0$ and $N(J)=0$ for $POS < J \leq 1000$. Thus

$$\text{time to next event} = -[\log(R_1)]/Q \quad .$$

(ii). Type of event : lines 34-53

The type of event is determined in an analagous manner to that described in section 7.2 for the two-colony model. An event occurs in colony I if

$$\sum_{K=25}^{24+I} N(I+K) < R_2[N(25)+\ldots+N(POS)] \leq \sum_{K=25}^{25+I} N(I+K) ,$$

and the type of event is

(a) a birth if $0 \leq R_3 < \lambda/(\lambda+\nu)$

(b) a migration one step to the right if $\lambda/(\lambda+\nu) \leq R_3 \leq 1$.

(iii). Output : lines 54-63

The sizes of the 25 leading colonies N(POS-24),...,N(POS) are printed each time the wavefront advances by one unit. A typical example ($\lambda=1$, $\nu_1=1$) of output is given in Table 9.1, though for reasons of space only the leading 6 colonies are shown. The first column (P) denotes the site of the wavefront, and the second column (T) denotes the time at which this site was first occupied.

The simulated population wave develops in a series of instantaneous jumps, and so any definition of a stochastic velocity or population waveform must involve some type of local 'smoothing' operation. The result of such a smoothing would itself be stochastic, but it should have (in some sense) a stationary distribution in the limit as T tends to infinity. A full-scale simulation of velocity and waveform distributions would be an interesting exercise (see Bailey (1967) and Mollison (1972a) for descriptions of simulated spatial stochastic epidemics), and work in the general area of simulated spatial stochastic processes certainly needs further development.

We may make some progress even with the simple set of results given in Table 9.1. For I have already shown (result (6.4)) that in the vicinity of the wavefront the asymptotic mean population size decreases geometrically at rate

(8.1) $$2\nu_1/[c_+ + (c_+^2+4\nu_1\nu_2)^{\frac{1}{2}}] \quad .$$

Table 9.1

A simulated wavefront showing the sizes of the 6 leading colonies immediately after the advance of the population wave: $\lambda=\nu_1=1$, $\mu=\nu_2=0$

P	T						
0	0.0						1
1	1.16					1	1
2	1.77				0	2	1
3	1.98			0	3	1	1
4	2.74		0	4	4	3	1
5	3.33	0	0	6	9	1	1
6	4.24	0	1	18	14	9	1
7	4.79	0	18	24	26	10	1
8	4.86	20	25	29	13	0	1
9	5.04	33	34	14	1	1	1
10	5.35	46	30	7	4	2	1
11	5.74	54	24	9	4	3	1
12	5.95	35	13	7	5	0	1
13	6.02	11	9	5	0	0	1
14	7.18	87	43	15	16	6	1
15	7.34	55	24	20	8	1	1
16	7.72	60	27	17	7	1	1
17	7.75	29	15	8	2	0	1
18	8.98	178	82	39	18	9	1
19	9.01	86	44	17	10	0	1
20	9.50	124	54	23	6	2	1
21	9.62	62	39	8	4	1	1
22	9.89	59	30	9	2	3	1
23	10.22	58	23	8	2	4	1
24	10.54	51	16	8	8	3	1
25	11.13	79	29	24	14	4	1

P - Position of wavefront

T - Time of wavefront advance

Now it follows from Figure 9.2 that when $\lambda=1$, $\mu=\nu_2=0$ and $\nu_1=1$ the velocity c_+ is equal to e, whence (8.1) becomes e^{-1}. Thus the expected values which correspond to $P=\infty$ in Table 9.1 are

$$e^5, \quad e^4, \quad e^3, \quad e^2, \quad e, \quad 1$$

i.e., to one decimal place,

(8.2) 148.4, 54.6, 20.1, 7.4, 2.7, 1 .

The corresponding averages of the last 10 values (P=16,...,25) in each column are

(8.3) 78.6, 35.9, 16.1, 7.3, 2.7, 1 .

The comparison between the theoretical values (8.2) and the simulated values (8.3) is encouraging, especially in view of the fact that the expected numbers (8.2) are asymptotic values. Note that when

(i) P=16 the mode is only 9 units from the wavefront

(ii) P=25 the mode is only 12 units from the wavefront,

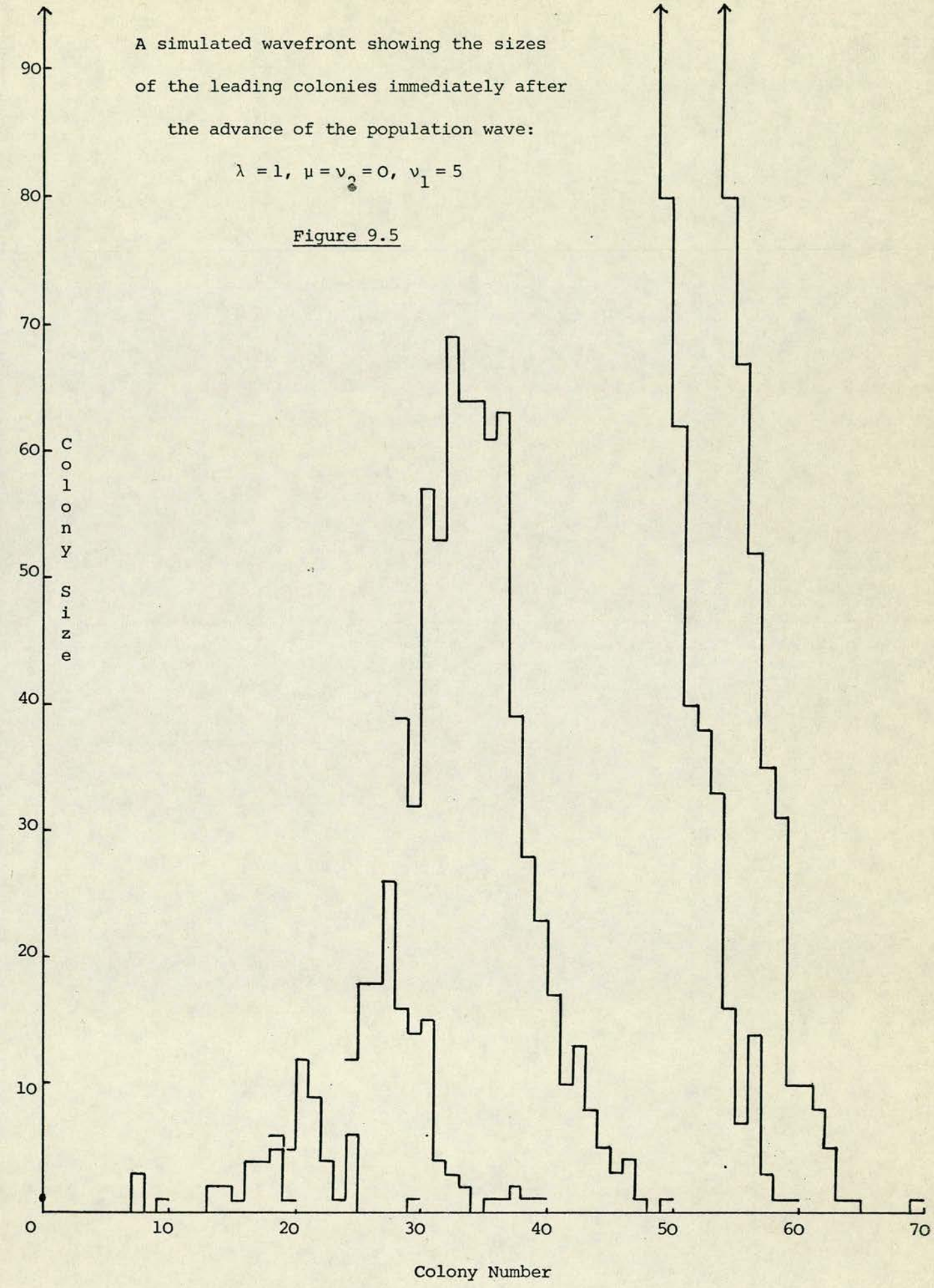
and so it is hardly reasonable to say that a colony which is 5 units from the wavefront is 'in the vicinity of the wavefront'. It is not unnatural that the first 3 numbers in (8.3) are lower than the corresponding values in (8.2), as result (8.1) will not apply to colonies this far from the 'wavefront'. The simulation clearly needs to be greatly extended before any 'statistical comparison' can be made, but unfortunately the computing time rises exponentially fast with P. This is clearly a situation that calls for ingenuity in simulation.

The results in Table 9.1 also support equation (3.20) which yields the (deterministic) velocity of propagation. A formal regression analysis of P against T over the range P=14,...,25

A simulated wavefront showing the sizes of the leading colonies immediately after the advance of the population wave:

$\lambda = 1,\ \mu = \nu_2 = 0,\ \nu_1 = 5$

Figure 9.5



yields a gradient of

2.7028

which compares very favourably with the value predicted from (3.20), namely

$$e = 2.7183 \quad \text{(to 4 dec. places)}.$$

When ν_1/λ is substantially larger than the value 1 considered above, migrations will predominate over births and the wavefront will be dominated by what is essentially a superposition of independent simple random walks. Thus there may exist several 'outliers' in front of the substantial mass of the wavefront. Figure 9.5 shows a realization of the build-up of a wavefront corresponding to the parameter values $\lambda=1$, $\mu=\nu_2=0$ and $\nu_1=5$ over the first 70 advances of the wave. The profile becomes extremely steep when $P=60$, and by $P=70$ one individual has already advanced 4 clear units ahead of the main wavefront.

CHAPTER 10

NON-NEAREST-NEIGHBOUR MIGRATION

10.1 Introduction

The theory developed in Chapter 8 is based on the assumption that an individual leaving colony i must immediately enter either colony $i-1$ or colony $i+1$. We shall now drop this assumption and consider instead the general migration distribution given by (9.2.1), namely

$$\Pr[\text{an individual in colony } j \text{ migrates to colony } i+j \text{ in } (t,t+\delta t)] = \nu h_i \delta t + o(\delta t) \qquad (-\infty < i,j < \infty)$$

where

$$\sum_{i=-\infty}^{\infty} h_i = 1 \quad ; \quad h_0 = 0 \ , \tag{9.2.2}$$

whence it follows (section 9.2) that

$$\sum_{i=-\infty}^{\infty} m_i(t) z^i = e^{(\lambda-\mu-\nu)t} \exp\{\nu t \sum_{j=-\infty}^{\infty} h_j z^j\} \ . \tag{9.2.6}$$

In this chapter expression (9.2.6) will be developed in two special cases. In the first $\{h_i\}$ is given by a geometric distribution, and in the second it assumes a more general form which is related to the spatial epidemic process studied by Mollison (1972a). Note that if the $\{h_i\}$ may be reasonably approximated by a normal distribution with a small mean and variance then the diffusion result (9.7.20) applies.

10.2 First-Order Moments for Geometric Migration Rates

Suppose

$$h_r = \begin{cases} \alpha^{r-1}(1-\alpha) & : \quad r=1,2,3,\ldots \\ 0 & : \quad r=0,-1,-2,\ldots \end{cases},$$

where $0<\alpha<1$, so that individuals may migrate only to the right and the probability of migration from colony i to colony $i+r$ decreases geometrically at rate α as r increases. Clearly

$$\sum_{r=-\infty}^{\infty} h_r = 1$$

as required, whilst for $|\alpha z| < 1$

$$\sum_{r=-\infty}^{\infty} h_r z^r = (1-\alpha)\sum_{r=1}^{\infty} \alpha^{r-1} z^r = z(1-\alpha)/(1-\alpha z) \quad . \tag{2.1}$$

On substituting (2.1) into (9.2.6) we obtain

$$\begin{aligned} \sum_{n=-\infty}^{\infty} m_n(t) z^n &= e^{(\lambda-\mu-\nu)t} \exp\{\nu t z(1-\alpha)(1-\alpha z)^{-1}\} \\ &= e^{(\lambda-\mu-\nu)t} \sum_{r=0}^{\infty} [\nu(1-\alpha)tz]^r (1-\alpha z)^{-r}/r! \\ &= e^{(\lambda-\mu-\nu)t}\left[1 + \sum_{r=1}^{\infty} [\nu(1-\alpha)tz]^r \sum_{s=0}^{\infty} \binom{r+s-1}{s} (\alpha z)^s/r!\right] \quad . \end{aligned} \tag{2.2}$$

The $m_n(t)$ $(n=0,1,2,\ldots)$ may be derived as the coefficients of z^n in (2.2), namely

$$m_0(t) = e^{(\lambda-\mu-\nu)t} \tag{2.3}$$

and

$$m_n(t) = e^{(\lambda-\mu-\nu)t} \alpha^n \sum_{r=1}^{n} \binom{n-1}{n-r} [\nu t(1-\alpha)/\alpha]^r/r! \qquad (n=1,2,\ldots) \quad . \tag{2.4}$$

The velocity of propagation corresponding to the mean values (2.3) and (2.4) has yet to be determined.

10.3 More Extreme Migration Rates

10.3.i Generating function for first-order moments

Mollison (1972a) simulated the spatial propagation of infection for simple epidemic models. In these models the rate of infection of susceptibles at a point s at time t is proportional to the product of the number of susceptibles at s with an average value of the numbers of infectious individuals at all points, weighted according to their distances from s. Thus the total influence of all the infectives is taken to be

$$\int_{\text{space}} Y(s-r)\,dV(r) \quad ,$$

where $Y(x)$ denotes the number of infectives at the point x and $V(r)$ is a weighting function.

He simulates the development of the wavefront for this spatial process for various forms of the weighting function $V(x)$. When $V(x)$ is exponentially bounded, so that

$$(3.1) \qquad \int_{\text{space}} e^{kx}\,dV(x) < \infty \qquad \text{for some constant } k > 0 \ ,$$

the wavefront is observed to advance at a steady rate. When condition (3.1) is not satisfied and $V(x)$ is of 'just infinite variance' ($\int x^{2-\varepsilon}\,dV(x)$ converges for arbitrarily small $\varepsilon > 0$ but diverges for $\varepsilon \leq 0$) the wavefront progresses in wilder and wilder leaps forward. Mollison's (1972a) paper contains several interesting diagrams which show the development of such simulated wavefronts.

A particularly interesting case is an intermediate one, when $V(x)$ is of finite variance but is not exponentially bounded. Here the simulated wavefronts develop in a mixture of steady progress and

great leaps forward and this behaviour could not be forecast by local approximation equations. Mollison remarks that if one could show that the distributions of light windborne objects (such as certain kinds of germs and plant seeds) are of this type, new light might be thrown on a number of problems which involve geographical spread. For example, it might explain why outbreaks of epidemics or mutant species sometimes appear to have several origins. For further references see Chamberlain (1970); Davies, Lewis and Randall (1968), and Norris and Harper (1970). Tinline (1970) considers the related problem of the 1967-1968 outbreak of foot and mouth disease.

In order to simulate the propagation of a spatial epidemic which has a weighting function that does not satisfy condition (3.1), Mollison (1972a) let $V(x)$ be of the form

$$k_r\Big[\prod_{u=1}^{r}(|x|+u)\Big]^{-1} \qquad (-\infty < x < \infty) \tag{3.2}$$

for r equal to 3 and 4. Here k_r is the appropriate normalizing constant. Although no analytic results are available for his epidemic process, we may develop a representation for the mean values $m_n(t)$ $(n=0,1,2,\ldots)$ for the corresponding migration process with one-way migration.

Let us write the analogue to (3.2) in the form

$$h_i = k_r\Big[\prod_{u=0}^{r}(i+u)\Big]^{-1} \qquad (i=1,2,\ldots\ ;\ r \geq 1)\ . \tag{3.3}$$

Then the p.g.f.

$$S_r(z) \equiv \sum_{n=-\infty}^{\infty} h_n z^n = k_r \sum_{n=1}^{\infty} \frac{z^n}{n(n+1)\dots(n+r)}$$

$$= k_r \sum_{n=0}^{\infty} \frac{z^{n+1}\Gamma(1+n)\Gamma(1+n)}{n!\Gamma(r+2+n)} \quad , \tag{3.4}$$

where $\Gamma(x)$ denotes the Gamma function. But the Gauss hypergeometric series is defined by

$$ {}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{z^n\Gamma(a+n)\Gamma(b+n)}{n!\Gamma(c+n)} \qquad (|z| \leq 1) \tag{3.5}$$

(Abramowitz and Stegun (1965), result (15.1.1)), and a comparison of expressions (3.4) and (3.5) yields

$$S_r(z) = [zk_r/(r+1)!]\, {}_2F_1(1,1;r+2;z) \quad . \tag{3.6}$$

To determine the normalizing constant k_r, we place $z=1$ in expression (3.5) and use the results that when $z=1$

$$S_r(1) = 1$$

and

$$ {}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \qquad (c \neq 0,-1,-2,\dots \; ; \; \Re(c-a-b) > 0) \tag{3.7}$$

(Abramowitz and Stegun (1965), result (15.1.20)). This yields

$$k_r = r.r! \tag{3.8}$$

whence (3.6) becomes

$$S_r(z) \equiv \sum_{n=-\infty}^{\infty} h_n z^n = [zr/(r+1)]\, {}_2F_1(1,1;r+2;z) \quad . \tag{3.9}$$

Substituting (3.9) into (9.2.6) we obtain the solution

$$\sum_{n=0}^{\infty} m_n(t) z^n = e^{(\lambda-\mu-\nu)t} \exp\{[\nu trz/(r+1)]\, {}_2F_1(1,1;r+2;z)\} \quad . \tag{3.10}$$

The exponent of the exponential in expression (3.10) is a power series in z, and so expressions for the $m_n(t)$ will be difficult to derive. However, an alternative representation may be obtained for this generating function which appears to have a more promising form.

On splitting h_n into partial fractions we may write

$$S_r(z) = r.r! \sum_{n=1}^{\infty} \frac{z^n}{n(n+1)\dots(n+r)}$$

$$= r.r! \sum_{n=1}^{\infty} \{z^n \frac{[r!]^{-1}}{n} + \dots + \frac{(-1)^i [i!(r-i)!]^{-1}}{n+i} + \dots + \frac{(-1)^r [r!]^{-1}}{n+r}\}$$

$$= r.r! \sum_{i=0}^{r} \frac{(-1)^i}{z^i i!(r-i)!} [-\log(1-z) - \sum_{s=1}^{i} (z^s/s)]$$

[we denote $\sum_{s=1}^{i} (z^s/s) \equiv 0$ when $i=0$]

$$(3.11) \qquad = -r(1-z^{-1})^r \log(1-z) - r \sum_{i=1}^{r} \sum_{s=1}^{i} s^{-1} \binom{r}{i} (-1)^i z^{s-i} \quad .$$

Combining (3.11) and (9.2.6) we obtain the solution

$$(3.12) \quad \sum_{n=0}^{\infty} m_n(t) z^n = e^{(\lambda-\mu-\nu)t} (1-z)^{-r\nu t} (1-z^{-1})^r \times \exp\{-r\nu t \sum_{i=1}^{r} \sum_{s=1}^{i} s^{-1} \binom{r}{i} (-1)^i z^{s-i}\} \quad .$$

Ideally we would like to obtain the $m_n(t)$ from the coefficients of the z^n in the expansion of either expressions (3.10) or (3.12), and from these mean values develop a family of velocities corresponding to $r = 1,2,3,\dots$. Moments of the distribution (3.3) for $\{h_i\}$ are finite up to order $r-1$ whilst all moments of higher order are infinite, and it would be interesting to compare the velocities for different values of r . If the same qualitative results hold for our spatial migration

process as for Mollison's epidemic process, the velocities of greatest interest should correspond to $r=2$ and $r=3$. A simulation study of the velocity of propagation of the wavefront has not yet been performed, and so at the present time we may only make the following conjectures.

(i) When $r=2$ the mean is finite and the variance is 'just infinite' ($\sum_{i=0}^{\infty} \frac{i^{2-\varepsilon}}{i(i+1)(i+2)}$ converges for arbitrarily small $\varepsilon > 0$ and diverges for $\varepsilon \leq 0$) - the wavefront will progress by greater and greater leaps forward as individuals in the vicinity of the wavefront migrate by larger and larger jumps because of the infinite variance of $\{h_2\}$.

(ii) When $r=3$ the mean and variance are both finite - the wavefront will advance in a mixture of steady progress and leaps forward, but the finite variance will ensure that the leaps will be sufficiently small for the population to 'catch up on itself'.

In the above discussion it must not be forgotten that observed or simulated wavefronts refer to a stochastic development, whilst theoretical velocities, which are calculated from mean values, refer to a deterministic development. For small values of r the theoretical and observed velocities will therefore almost certainly have little in common because of the 'leaping behaviour' described above. Unfortunately the only value of r for which we have isolated the $m_n(t)$ from their generating function is $r=1$, and as well as this being the smallest value of r allowed both the

mean and the variance of $\{h_1\}$ are infinite! However, the derivation of the $m_n(t)$ for $r=1$ is presented below as it demonstrates the type of solution we might expect for higher values of r.

10.3.ii First-order moments ($r=1$)

Placing $r=1$ in expression (3.12) we have

$$\sum_{n=0}^{\infty} m_n(t) z^n = e^{(\lambda-\mu)t}(1-z)^{-\nu t(1-z^{-1})} \quad . \tag{3.13}$$

Now

$$(x)_n \equiv x(x-1)\ldots(x-n+1) = \sum_{k=0}^{n} s(n,k) x^k \quad ,$$

where the $s(n,k)$ are Stirling Numbers of the First Kind (Riordan (1958), p.33). Thus

$$(1-z)^{\nu t z^{-1}} = 1 + \sum_{n=1}^{\infty} \frac{(-z)^n}{n!} (\nu t z^{-1})(\nu t z^{-1}-1)\ldots(\nu t z^{-1}-n+1)$$

$$= 1 + \sum_{n=1}^{\infty} \frac{(-z)^n}{n!} \sum_{k=0}^{n} s(n,k)(\nu t z^{-1})^k$$

whence

$$\text{coeff. of } z^r \equiv Q_r = (-1)^r \sum_{k=0}^{\infty} \frac{(-\nu t)^k}{(r+k)!} s(r+k,k) \quad . \tag{3.14}$$

But

$$(1-z)^{-\nu t} = 1 + \sum_{k=1}^{\infty} \frac{z^k}{k!} (\nu t)(\nu t+1)\ldots(\nu t+k-1) \quad , \tag{3.15}$$

and on combining (3.14) and (3.15) we see that

$$\text{coeff. of } z^n \text{ in } (1-z)^{-\nu t(1-z^{-1})} = \sum_{r=0}^{n} \frac{Q_r}{(n-r)!} (\nu t)(\nu t+1)\ldots(\nu t+n-r-1).$$

Thus

$$(3.16)\quad m_n(t) = e^{(\lambda-\mu)t} \sum_{r=0}^{n} \frac{Q_r}{(n-r)!}(-1)^{n-r}(-\nu t)_{n-r} \qquad (n=0,1,2,\ldots)$$

where the $\{Q_r\}$ are defined by expression (3.14).

CHAPTER 11

THE SPATIAL DISTRIBUTION OF TRIBOLIUM CONFUSUM

11.1 Introduction

In previous chapters we have developed theoretical results for population models in which there is migration either between two colonies or between an infinite number of colonies situated at the integer points of a single co-ordinate axis. To conclude this thesis we shall present a spatial model in two-dimensions, and relate it to data on the spatial distribution of flour beetles (Tribolium) in a closed container.

The genus Tribolium consists of 26 different species. Two of these species, Tribolium Confusum and Tribolium Castaneum, spend their entire life-span in finely milled flour, and are ideally suited for experiments on population growth and competition. The flour is both their habitat and food supply. Counts of such beetle populations may be made very accurately, as the infested flour can be sieved to recover beetles at all stages of their development. Moreover, the flour acquires the temperature and moisture characteristics of its external climate, and hence the 'microclimate' within the flour may be accurately controlled. Although the Tribolium experiment which is described below may appear rather artificial, it is more 'natural' than might seem at first glance. Andres (1931) reports the existence of Tribolium Confusum in a Pharaonic Tomb (about 2500 B.C.), and so the species has had at least 4,500 years in which to adapt genetically to its flour environment.

The use of flour beetles as 'experimental insects' does unfortunately have certain disadvantages. The length of a generation from egg to egg at 29°C is roughly a month, and so an experiment can take several years. Moreover, it is essentially impossible to observe the beetles' behaviour in the flour. Adult beetles cannot be quickly sexed, nor assessed as to age, and so it has been impossible to construct accurate age-and-sex distributions for various sorts of populations. However, although Tribolium is far from ideal, it does appear to have as great a combination of attributes favourable for ecological population study as any species so far used in laboratory analysis of such problems.

11.2 Description of a Study of Tribolium Confusum

Neyman, Park and Scott (1956) describe an experiment which they performed to provide information about the spatial distribution of living and dead adult beetles of both sexes after the beetles have been allowed to travel for a considerable time.

A container was filled with fresh flour occupying the volume of a cube 10×10×10 inches. On the surface of the flour a total of 2257 adult Tribolium Confusum beetles was placed in a 1 to 1 sex ratio (the authors do not state whether the beetles were put in one spot or spread evenly over the surface). The container was kept in a dark incubator at an approximately constant temperature of 29°C and an approximately constant humidity of 70%. Also, in order to equalize the possible gradients of temperature within the incubator, the container was periodically rotated. After 4 months, at which time the population totalled 73,009 individuals excluding

eggs, the contents of the container were divided into 1000 cubes 1×1×1 inch and the contents of each cube were lifted separately and examined.

The data in Appendix C represent the second layer from the top of the experimental cube and show the distribution of adult males and females within the 100 constituent small cubes. The general character of this distribution in other layers is very similar. However, the average density of beetles shows a marked decrease from the top of the container downwards. [As the authors do not define density the number of male and female beetles in each of the 100 cubes is only known to within a multiplicative constant].

The distributions for males and females are somewhat different in that the density for the females falls steadily towards the centre of the square from a relatively low maximum at the corners, while the density for the males falls more sharply from a high maximum at the corners and is substantially constant in the central portion. However, the general feature of a gradual increase in density towards the edges and along the edges towards the corners is the same for both sexes.

11.3 Previous Work on the Data

A number of random walk models have been tried in an effort to obtain a limiting distribution which possesses the same characteristics as the empirical distribution described above. Some of these models have met with more success than others, and it is of interest to mention some of the results.

Broadbent and Kendall (1953) had some success in describing the two-dimensional motion of Trichostrongylus Retortaeformis by simple Brownian motion. Sherman investigated the consequences of such a model in the present context, and showed that a random walk over a square lattice with inelastic boundaries could produce a concentration of beetles at the boundaries and in the corners, but could not account for the gradual increase in density towards the edges. Sherman (1956) investigated the one-dimensional random walk and showed that distributions more like the one observed could be obtained by the use of a suitable boundary condition. This was that after striking the boundary, the beetle remains there a randomly distributed time, and is then placed instantaneously a finite distance within the region of motion. He did not, however, claim this boundary condition to be a reasonable explanation of the experimental results.

Cox and Smith (1957) assume that the motion of the beetle is described by a more complicated random walk whose steps are not infinitesimal, and may be correlated in direction and length (see, for example, the type of walk investigated by Daniels (1952)). On meeting the boundary the beetle remains there for a time-interval which has a certain frequency distribution, and it then pursues a path that is the reflexion in the boundary of the path that would have been followed in the absence of the boundary. The authors remark that this model leads to a limiting distribution which is uniform within the square boundary and is a line concentration on the boundary. Moreover, this result applies also

for motion within a circular boundary, and is probably true for a very wide class of boundaries. Cox and Smith then go on to obtain a distribution of beetles of the type observed, by the use of a more realistic boundary condition.

Their basic assumptions are that motion is two-dimensional and that within the flour a beetle follows a straight path. When a beetle meets the boundary it may (i) with probability p, return along its original path; or, (ii) with probability $1-p$, move a distance s (possibly zero) along the boundary before choosing, independently of its previous direction, a new direction of motion. The distance s and the angle θ that the new direction of motion makes to the normal to the boundary at the point of departure are assumed to have distribution functions $H(s)$ and $G(\theta)$, respectively. For simplicity they assume that the beetle moves with constant speed, although they remark that this assumption may be greatly generalized without affecting the final results.

Assume that motion takes place within a circle of unit radius and let $\bar{s}$ denote the expected value of s. Then Cox and Smith (1957) show that if $w(r)$ denotes the density of beetles per unit area as a function of r, and

$$G'(\theta) = a + b|\sin\theta| + c\sin^2\theta \qquad (a,b,c \text{ are constants}),$$

then

$$w(r) = \frac{aK(r) + \tfrac{1}{2}b\log[(1+r)/(1-r)] + c(K(r)-E(r))}{\pi[2a + b + (2c/3) + \bar{s}(1-p)]}$$

where

$$K(r) = \int_0^{\pi/2}(1-r^2\sin^2\theta)^{-\frac{1}{2}}d\theta \quad ; \quad E(r) = \int_0^{\pi/2}(1-r^2\sin^2\theta)^{\frac{1}{2}}d\theta$$

are complete elliptic integrals of the first and second kinds, respectively. In addition to this continuous distribution over the interior of the unit circle, there is a line density of probability on the circumference of amount per unit length

$$\frac{\bar{s}(1-p)/(2\pi)}{2a + b + (2c/3) + \bar{s}(1-p)} .$$

The general conclusions that they draw from these results are as follows.

(i) $b=c=0$: θ uniform - The beetles select paths which are randomly oriented, and the resulting density is nearly constant for $r < \frac{1}{2}$ and increases steadily to infinity as r tends to one.

(ii) $a=c=0$: $G'(\theta) \propto |\sin\theta|$ - There is a fairly strong tendency to select paths nearly tangential to the boundary, and the resulting distribution of beetles is naturally much more strongly concentrated at the larger values of r, rising from a zero density at the centre.

(iii) $a=b=0$: $G'(\theta) \propto |\sin^2\theta|$ - This is a much stronger version of (ii).

The required distribution of beetles is taken to be a weighted combination of (i)-(iii).

One of my colleagues, Mr M.F. Franklin, has shown that it is relatively easy to simulate such paths on a computer and to output the cumulative paths on an oscilloscope (this technique involves the use of a PDP-15 machine). Polaroid photographs of

the build-up of these paths provide a permanent record.

Cox and Smith tried to extend their results to the case of the square boundary, but unfortunately they were unable to solve the integral equation which determines the distribution of starting points along the edges. However, they were able to obtain an approximate empirical solution.

Now whilst it is extremely unlikely that a beetle will follow a straight path within the flour, the Brownian motion models may be criticized on the grounds that they imply that the path of a beetle is highly irregular; as though the beetles were constantly forgetting their direction. Moreover, such models imply that the beetles travel with infinite velocity! An alternative scheme may be considered whereby the velocity is a stochastic process with independent increments, but this approach also has to be abandoned. For the model implies a correlation between a beetle's position at any given time and its corresponding velocity, and this is something which is unacceptable intuitively and is contradicted by observations based on X-ray photographs (see Neyman et al., p.75).

To balance these two extremes, of complete irregularity and straight line paths, we shall develop a stepping-stone model which produces a distribution of the required shape. In this model the beetles migrate over a set of lattice points, and may change their direction of motion at each point. For example, the Tribolium data is in the form of 100 squares each of side 1 inch, so the beetles are considered as migrating over a square lattice of size 10×10.

The equations for this model were in effect developed by Usher and Williamson (1970), although their interpretation of the parameters in the equations is different from mine. They consider a one-dimensional discrete time model, with colonies situated at the integer points $i=0,\ldots,n$. At each time point a proportion of the beetles migrate to a nearest-neighbour and the remainder stay where they are; the birth and death rates of the 'movers' and 'stayers' are λ_2, μ_2 and λ_1, μ_1 , respectively.

Put $R_1=1+\lambda_1-\mu_1$, $R_2=1+\lambda_2-\mu_2$ and let the migration rate between any two neighbouring colonies be ν , where

$$0 \leq \lambda_1,\lambda_2 \quad ; \quad 0 \leq \mu_1,\mu_2 \leq 1 \quad ; \quad 0 \leq \nu \leq \tfrac{1}{2} \ .$$

Then $m_i(t)$, the mean number of beetles in colony i at time t , is given by the set of equations

$$\begin{aligned} m_0(t+1) &= (1-\nu)R_1 m_0(t) + \nu R_2 m_1(t) \\ m_i(t+1) &= (1-2\nu)R_1 m_i(t) + \nu R_2 (m_{i-1}(t)+m_{i+1}(t)) \qquad (i=1,\ldots,n-1) \\ m_n(t+1) &= (1-\nu)R_1 m_n(t) + \nu R_2 m_{n-1}(t) \ . \end{aligned}$$

(3.1)

Usher and Williamson write equations (3.1) in the equivalent vector-matrix form

$$\underline{m}(t+1) = M\underline{m}(t) \quad ,$$

and note that if the population eventually reaches a steady state, i.e. there is no proportional change in the elements of $\underline{m}(t)$, then

$$M\underline{m}(\infty) = \Lambda\underline{m}(\infty) \tag{3.2}$$

where Λ is the largest eigenvalue of M (Λ corresponds to the overall rate of growth). They rewrite equations (3.1) as

$$(3.3) \qquad \begin{aligned} m_1 &= \{[\Lambda - R_1(1-\nu)]/\nu R_2\} m_0 \\ m_i &= \{[\Lambda - R_1(1-2\nu)]/\nu R_2\} m_{i-1} - m_{i-2} \qquad (i=2,\ldots,n) , \end{aligned}$$

and show numerically that (3.3) yields a distribution of the required shape. They extend this model to two and three dimensions and apply it to the Tribolium data, obtaining the conclusion that 'beetles which move around are less successful breeders than those that stay put'. Although the model fits the female data quite well, the fit for the males is unfortunately not so good, as the theoretical density of males is too great towards the centre and is insufficient towards the corners.

11.4 A One-Dimensional Stepping-Stone Model with Nearest-Neighbour Migration

11.4.i First-order moments

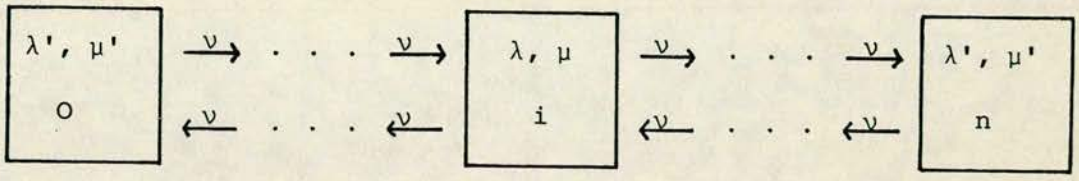
The model of Usher and Williamson (1970) assumes that beetles which move in a given time period have different birth and death rates to those which are stationary. We now propose an alternative model, namely that the birth and death rates within the region of motion are different from those on the boundary. Both models yield the same set of equations, although the interpretation of them is different, and so the following analysis also applies to the 'mover-stayer' situation.

Let the birth and death rates of colony i $(i=1,\ldots,n-1)$ be λ and μ, and those of colonies 0 and n be λ' and μ'. Denote the migration rate between nearest-neighbours by ν, with $m_i(t)$ $(i=0,\ldots,n\ ;\ t=0,1,2,\ldots)$ representing the number of

beetles in colony i at the discrete time point t (this situation is illustrated in figure 11.1). Then if we assume that their

Figure 11.1

One-dimensional stepping-stone model with different rates on the boundary



proportional values settle down to steady values for large enough t (say $t >$ some $T > 0$), i.e.

$$m_i(t+1) = \omega m_i(t) \qquad (i=0,\ldots,n \; ; \; t > T) \tag{4.1}$$

where ω is the overall rate of growth, we have the following set of equations:

$$\begin{aligned}
m_0(t+1) = \omega m_0(t) &= (1+\lambda'-\mu'-\nu)m_0(t) + \nu m_1(t) \\
m_i(t+1) = \omega m_i(t) &= (1+\lambda-\mu-2\nu)m_i(t) + \nu m_{i-1}(t) + \nu m_{i+1}(t) \quad (i=1,\ldots,n-1) \\
m_n(t+1) = \omega m_n(t) &= (1+\lambda'-\mu'-\nu)m_n(t) + \nu m_{n-1}(t) \quad .
\end{aligned} \tag{4.2}$$

Here

$$0 \leq \lambda, \lambda' \quad ; \quad 0 \leq \mu, \mu' \leq 1 \quad ; \quad 0 \leq \nu \leq \tfrac{1}{2}$$

with

$$\mu' + \nu \leq 1 \quad ; \quad \mu + 2\nu \leq 1 \quad ,$$

and, unlike the 'mover-stayer' model, a beetle is not allowed simultaneously to migrate and to give birth.

Let us assume that $t > T$. Then as we are interested in the relative and not the absolute values of the $m_i(t)$ we shall now

disregard time, so that the parameter t will not explicitly feature in the following results.

If we write

$$r = [\omega-(1+\lambda-\mu-2\nu)]/(2\nu)$$
$$r' = [\omega-(1+\lambda'-\mu'-\nu)]/\nu \tag{4.3}$$

equations (4.2) become

$$r'm_0 = m_1 \tag{4.4}$$
$$2rm_i = m_{i-1} + m_{i+1} \qquad (i=1,\ldots,n-1) \tag{4.5}$$
$$r'm_n = m_{n-1} \quad . \tag{4.6}$$

[Note that equations (3.3) and (4.4)-(4.6) are identical if

$$r = [\Lambda-R_1(1-2\nu)]/(2\nu R_2)$$

and $$r' = [\Lambda-R_1(1-\nu)]/(\nu R_2) \quad] \ .$$

Put

$$r = \cosh\theta \quad . \tag{4.7}$$

The difference equation (4.5) has the associated auxiliary equation

$$z^2 - 2z\cosh\theta + 1 = 0$$

which has roots e^{θ}, $e^{-\theta}$. Thus (4.5) has the general solution

$$m_i = Ae^{i\theta} + Be^{-i\theta} \tag{4.8}$$

where A and B are constants. For reasons of symmetry we require

$$m_0 = m_n$$

whence

$$A + B = Ae^{n\theta} + Be^{-n\theta} \quad . \tag{4.9}$$

From (4.9) we get

$$B = Ae^{n\theta}$$

and so from (4.8) we have

$$m_i = k.\cosh[\tfrac{1}{2}(n-2i)\theta] \qquad (i=0,\ldots,n) \tag{4.10}$$

where k is a suitably chosen normalizing constant.

Expression (4.10) is very closely related to the equation of the common catenary,

$$y = c.\cosh(x/c) \quad ,$$

which is the curve describing the manner in which a uniform chain or a 'perfectively flexible' string hangs when freely suspended from two fixed points. This relationship enables θ to be estimated 'by eye'! For if observed values of $\{m_i\}$ are plotted on graph paper, which is then supported vertically, a uniform piece of string may be held next to the paper and moved into a position judged to give the 'best fit'. The parameter θ in (4.10) may then be computed from knowledge of the length of the string between any two given points (see, for example, Ramsey (1960), p.238).

Whilst the parameter θ may be estimated numerically from (4.10) by least squares techniques, an extremely simple estimator of θ may be obtained by observing the 'drop' between the two 'end' observations m_0, m_n and estimating the 'central' value $m_{n/2}$ (an interpolation approximation is necessary if n is odd). Without loss of generality we may define

$$m_0 = m_n = 1 \quad .$$

At $i=0$ or $i=n$ expression (4.10) gives

$$1 = k.\cosh(\tfrac{1}{2}n\theta)$$

whence

$$m_i = \frac{\cosh[\tfrac{1}{2}(n-2i)\theta]}{\cosh(\tfrac{1}{2}n\theta)} \qquad (i=0,\ldots,n) \quad . \tag{4.11}$$

If we put $i=\frac{1}{2}n$ (not necessarily integer) we have

$$m_{n/2} = [\cosh(\tfrac{1}{2}n\theta)]^{-1}$$

which yields the estimator

$$\hat{\theta} = (2/n)\cosh^{-1}(m_{n/2}^{-1}) \quad . \tag{4.12}$$

The parameter θ may also be determined analytically in terms of the solution of a certain equation. For it follows from (4.3) and (4.7) that

$$2\nu\cosh\theta = \omega - (1+\lambda-\mu-2\nu) \quad , \tag{4.13}$$

whilst a combination of (4.10) with either (4.4) or (4.6) gives

$$r'\cosh(\tfrac{1}{2}n\theta) = \cosh[\tfrac{1}{2}(n-2)\theta] \quad . \tag{4.14}$$

Eliminating ω between (4.3) and (4.13) we get

$$r' = 2\cosh\theta + (1/\nu)(\lambda-\mu-\lambda'+\mu'-\nu) \quad ,$$

whence (4.14) becomes, after a little simplification,

$$\cosh[\tfrac{1}{2}(n+2)\theta] = [1 - (1/\nu)(\lambda-\mu-\lambda'+\mu')]\cosh(\tfrac{1}{2}n\theta) \quad . \tag{4.15}$$

If the parameter values λ, λ', μ, μ' and ν are known, equation (4.15) may be solved numerically for θ.

11.4.ii The characteristic equation

For any given set of parameter values the maximal eigenvalue ω may be computed from (4.13) and (4.15). However, to determine the remaining eigenvalues it is necessary to revert to the original equations (4.2). The eigenvalue of second largest magnitude is of particular importance as it determines the rate of convergence to the 'steady-state' solution.

If we write

$$s = (1+\lambda-\mu-2\nu)/(2\nu) \quad ; \quad s' = (1+\lambda'-\mu'-\nu)/\nu$$

equations (4.2) become

$$\begin{aligned} m_0(t+1) &= \nu s' m_0(t) + \nu m_1(t) \\ m_i(t+1) &= 2\nu s m_i(t) + \nu m_{i-1}(t) + \nu m_{i+1}(t) \qquad (i=1,\ldots,n-1) \\ m_n(t+1) &= \nu s' m_n(t) + \nu m_{n-1}(t) \quad . \end{aligned} \tag{4.16}$$

Hence the matrix of transition rates is given by

$$Q \equiv \begin{bmatrix} \nu s' & \nu & 0 & & \\ \nu & 2\nu s & \nu & & \\ 0 & \nu & 2\nu s & & \\ & & & \ddots & \\ & & & & \nu s' \end{bmatrix} ,$$

and the characteristic equation

$$|Q - qI| = 0$$

may be written as

$$\begin{vmatrix} (s'-q/\nu) & 1 & 0 & & \\ 1 & (2s-q/\nu) & 1 & & \\ 0 & 1 & (2s-q/\nu) & & \\ & & & \ddots & \\ & & & & (s'-q/\nu) \end{vmatrix} = 0 \quad . \tag{4.17}$$

To solve this determinantal equation we need two results.

Firstly, the $m \times m$ determinant

$$(4.18)\qquad D_m \equiv \begin{vmatrix} a & 1 & 0 & & \\ 1 & a & 1 & & \\ 0 & 1 & a & & \\ & & & \ddots & \\ & & & & a \end{vmatrix} = \sinh[(m+1)\psi]/\sinh(\psi)$$

where $a=2\cosh(\psi)$. We obtain this expression by expanding the top row of D_m and solving the resulting difference equation. Secondly, the $(n+1)\times(n+1)$ determinant E_{n+1} may be written as

$$(4.19)\qquad E_{n+1} \equiv \begin{vmatrix} b & 1 & 0 & & & \\ 1 & a & 1 & & & \\ 0 & 1 & a & & & \\ & & & \ddots & & \\ & & & & a & 1 \\ & & & & 1 & b \end{vmatrix} = b^2 D_{n-1} - 2bD_{n-2} + D_{n-3} \quad .$$

Here we expand the top and bottom rows of E_{n+1} .

A comparison of (4.17), (4.18) and (4.19) with

$$b = (s'-q/\nu) \quad ; \quad a = 2\cosh(\psi) = (2s-q/\nu)$$

yields the equation

$$(4.20)\qquad \{(s'-q/\nu)^2\sinh(n\psi) - 2(s'-q/\nu)\sinh[(n-1)\psi] + \sinh[(n-2)\psi]\} \div \sinh(\psi) = 0 \quad .$$

The eigenvalues $\{q_i\}$ $(i=0,\ldots,n)$ are given by the roots of equation (4.20). In particular, it is easily verified that if $q=\omega$ equation (4.20) may be reduced to (4.14).

This argument may be extended a little. For if we assume that $\sinh(\psi)\neq 0$ then we may reduce equation (4.20) to

$$(s'-2s)^2(e^{(2n+2)\psi}-e^{2\psi}) + 2(s'-2s)(e^{(2n+3)\psi}-e^{\psi}) + (e^{(2n+4)\psi}-1) = 0$$

which is a polynomial of order $2(n+2)$ in e^{ψ}. There will be $2(n+2)$ roots which are conjugate in pairs. One of these pairs will correspond to $\psi = 0$ which may be neglected because it conflicts with our requirement that $\sinh(\psi) \neq 0$. There remain $2(n+1)$ conjugate roots of which we need those given by $\psi > 0$. Let $\eta_i = \exp(\psi_i)$ $(i=0,\ldots,n)$ be these roots. Then as

$$2\cosh(\psi) = 2s - q/\nu$$

we have

$$2\cosh[\log(\eta_i)] = 2s - q_i/\nu$$

and so

$$q_i = 2\nu\{s - \cosh[\log(\eta_i)]\} \qquad (i=0,\ldots,n) \quad . \tag{4.21}$$

11.5 A Two-Dimensional Stepping-Stone Model with Nearest-Neighbour Migration

Consider a square lattice defined by the set of points (i,j) where $i,j = 0,\ldots,n$. Let the birth and death rates of colonies

(i) in the interior be λ and μ

(ii) on the sides be λ' and μ'

(iii) in the corners be λ'' and μ''.

Denote the migration rate between nearest-neighbours by ν, and let $m_{ij}(t)$ $(i,j=0,\ldots,n\ ;\ t=0,1,2,\ldots)$ represent the number of beetles in colony (i,j) at the discrete time point t. This situation is illustrated in figure 11.2.

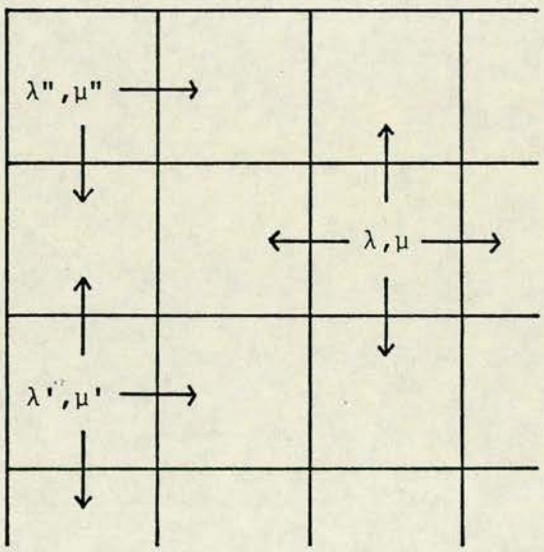
Put

$$H_{ij}(t) = m_{i+1,j}(t) + m_{i-1,j}(t) + m_{i,j+1}(t) + m_{i,j-1}(t) \quad (0 \leq i,j \leq n)$$

and define $m_{ij}(t) \equiv 0$ if (i,j) is not a lattice point (i.e. if either $i,j < 0$ or $i,j > n$) .

Figure 11.2

Two-dimensional stepping-stone model with different rates on the boundary



Then if we assume that the proportional values of the $m_{ij}(t)$ settle down to steady values for large enough t (say $t >$ some $T_1 > 0$), i.e.

$$m_{ij}(t+1) = \omega m_{ij}(t) \qquad (i,j = 0,\ldots,n \; ; \; t > T_1) \tag{5.1}$$

where ω is the overall rate of growth, we see from figure 11.2 that we have the following set of equations for $i,j = 0,\ldots,n$.

$$(5.2)\quad m_{ij}(t+1) = \omega m_{ij}(t) = \begin{cases} (1+\lambda-\mu-4\nu)m_{ij}(t)+\nu H_{ij}(t) & : \text{INTERIOR} \\ \qquad\qquad (\text{neither } i \text{ nor } j = 0 \text{ or } n) & \\ (1+\lambda'-\mu'-3\nu)m_{ij}(t)+\nu H_{ij}(t) & : \text{SIDES} \\ \qquad\qquad (i=0,n \text{ or } j=0,n \text{ but not both}) & \\ (1+\lambda"-\mu"-2\nu)m_{ij}(t)+\nu H_{ij}(t) & : \text{CORNERS} \\ \qquad\qquad (i=0,n \text{ and } j=0,n). & \end{cases}$$

Here

$$0 \le \lambda,\lambda',\lambda" \quad ; \quad 0 \le \mu,\mu',\mu" \le 1 \quad ; \quad 0 \le \nu \le \tfrac{1}{4}$$

and

$$\mu"+2\nu \le 1 \quad ; \quad \mu'+3\nu \le 1 \quad ; \quad \mu+4\nu \le 1 \quad .$$

If we write

$$(5.3)\qquad a = \omega - 1$$

equations (5.2) reduce to

$$(5.4)\quad H_{ij}(t)/m_{ij}(t) = \begin{cases} K \equiv \nu^{-1}(a-\lambda+\mu) + 4 & : \text{INTERIOR} \\ K' \equiv \nu^{-1}(a-\lambda'+\mu') + 3 & : \text{SIDES} \\ K" \equiv \nu^{-1}(a-\lambda"+\mu") + 2 & : \text{CORNERS} . \end{cases}$$

Let us assume that $t > T_1$. Then as we are interested in the relative and not the absolute values of the $m_{ij}(t)$ we shall now disregard time (as we did for the one-dimensional model in the previous section).

In view of the nature of the solution (4.10) for the one-dimensional model, let us consider the trial solution

$$(5.5)\qquad m_{ij} = c.\cosh[(n-2i)\theta]\cosh[(n-2j)\theta] \qquad (i,j = 0,\ldots,n)$$

where c is a suitably chosen normalizing constant. If we substitute (5.5) into the 'interior' equation (5.4) we obtain

$$K.\cosh[(n-2i)\theta]\cosh[(n-2j)\theta]$$

$$= \cosh[(n-2j)\theta]\{\cosh[(n-2i-2)\theta] + \cosh[(n-2i+2)\theta]\}$$

$$+ \cosh[(n-2i)\theta]\{\cosh[(n-2j-2)\theta] + \cosh[(n-2j+2)\theta]\}$$

$$= 4\cosh(2\theta)\cosh[(n-2i)\theta]\cosh[(n-2j)\theta] \quad .$$

Hence for (5.5) to be a solution we require that

$$(5.6) \qquad K \equiv (\omega-1-\lambda+\mu+4\nu)/\nu = 4\cosh(2\theta) \quad .$$

Expression (5.6) relates θ to the overall growth rate ω .

To evaluate θ let us use the 'corner' equation (5.2), which for reasons of symmetry may be written as

$$(5.7) \qquad \omega m_{oo} = (1+\lambda"-\mu"-2\nu)m_{oo} + 2\nu m_{o1} \quad .$$

If we put the trial solution (5.5) into (5.7) we have

$$(5.8) \qquad (\omega-1-\lambda"+\mu"+2\nu)\cosh^2(n\theta) = 2\nu\cosh(n\theta)\cosh[(n-2)\theta] \quad ,$$

and substitution of ω from (5.6) into this expression yields the equation

$$(5.9) \qquad [(\lambda-\mu)-(\lambda"-\mu")-2\nu+4\nu\cosh(2\theta)] = 2\nu\cosh[(n-2)\theta]/\cosh(n\theta) \quad .$$

For a given set of parameter values λ, $\lambda"$, μ, $\mu"$ and ν equation (5.9) determines the value of θ and hence the growth rate ω . However, we have yet to consider the 'side' equation (5.2), which for reasons of symmetry may be written as

$$(5.10) \qquad \omega m_{io} = (1+\lambda'-\mu'-3\nu)m_{io} + \nu(m_{i-1,o} + m_{i+1,o} + m_{i,1}) \quad .$$

Substituting the trial solution (5.5) into (5.10) we have

$$(5.11) \qquad (\omega-1-\lambda'+\mu'+3\nu) = \nu\{2\cosh(2\theta) + \cosh[(n-2)\theta]/\cosh(n\theta)\} \quad ,$$

and so once the parameters λ, $\lambda"$, μ, $\mu"$ and ν are specified we are completely restricted in our choice of $\lambda'-\mu'$. Note that equation (5.11) does not contain any terms which involve the

location parameters i and j, and so $\lambda' - \mu'$ is invariant over all side colonies ($i=0,n$ or $j=0,n$ but not both) as required.

The solution (5.5) possesses the full symmetry properties of the square, and in order to compare it with the Tribolium data in Appendix C I have arranged for the latter to possess these same symmetry properties by computing the averages

$$n_{ii} = (1/4)(\hat{m}_{i,i} + \hat{m}_{i,9-i} + \hat{m}_{9-i,i} + \hat{m}_{9-i,9-i}) \qquad (0 \leq i \leq 4)$$

$$n_{ij} = (1/8)(\hat{m}_{i,j} + \hat{m}_{i,9-j} + \hat{m}_{9-i,j} + \hat{m}_{9-i,9-j} + \hat{m}_{j,i} + \hat{m}_{j,9-i} + \hat{m}_{9-j,i} + \hat{m}_{9-j,9-i}) \qquad (0 \leq i < j \leq 4)$$

for both the female and the male populations. Here $\hat{m}_{ij}$ $(i,j=0,\ldots,9)$ denotes the entry in the i^{th} row and the j^{th} column of the 10×10 data matrix in Appendix C (the upper figure in each cell is proportional to the density of females, the lower figure is proportional to the density of males). Tables 11.1a and 11.2a show the results of applying the transformation

$$n_{ij}/n_{44} \qquad (0 \leq i,j \leq 4)$$

to the female and male data, respectively. Whilst the density of the females falls steadily towards the centre of the square from a relatively low maximum at the corners, the density of the males falls more sharply from a high maximum at the corners. Moreover, the male density has a local maximum at the centre, which the distribution (5.5) cannot possibly emulate. [Note that a similar table in Cox and Smith (1957) on p.328, given to an accuracy of one decimal place, contains several errors].

Table 11.1a

Transformed female Tribolium data

(for transformation see text)

0	1	2	3	4	i / j
5.64	4.37	4.13	3.71	4.01	0
	2.89	2.50	2.15	2.02	1
		1.64	1.34	1.27	2
			1.57	1.10	3
				1.00	4

Table 11.1b

Values of the function $\ell_{ij} = \cosh[(9-2i)\hat{\theta}]\cosh[(9-2j)\hat{\theta}]/\cosh^2(\hat{\theta})$:

$\hat{\theta} = 0.17$

0	1	2	3	4	i / j
5.68	4.22	3.25	2.66	2.38	0
	3.13	2.41	1.98	1.77	1
		1.86	1.52	1.36	2
			1.25	1.12	3
				1.00	4

Table 11.2a

Transformed male Tribolium data

(for transformation see text)

0	1	2	3	4	i / j
13.31	8.15	6.15	5.39	6.26	0
	3.15	1.77	1.43	1.64	1
		0.92	0.97	0.88	2
			0.57	0.80	3
				1.00	4

Table 11.2b

Values of the function $\ell_{ij} = \cosh[(9-2i)\hat{\theta}]\cosh[(9-2j)\hat{\theta}]/\cosh^2(\hat{\theta})$:

$\hat{\theta} = 0.22$

0	1	2	3	4	i / j
12.98	8.58	5.87	4.31	3.60	0
	5.67	3.88	2.85	2.38	1
		2.65	1.95	1.63	2
			1.43	1.20	3
				1.00	4

To illustrate the differences in shape between the solution (5.5) and the Tribolium data, I have computed the function

$$\ell_{ij} \equiv \cosh[(9-2i)\theta]\cosh[(9-2j)\theta]/\cosh^2(\theta) \qquad (0 \leq i \leq j \leq 4)$$

and presented the values for the females and males in Tables 11.1b and 11.2b, respectively. Clearly $\ell_{44}=1$, and I have estimated the parameter θ by solving the equations

$$\ell_{00} = 5.64 \text{ (females)} \quad ; \quad \ell_{00} = 13.31 \text{ (males)} .$$

This is easily done to an accuracy of two decimal places by an inspection of cosh tables, which gives

$$\hat{\theta} = 0.17 \text{ (females)} \quad ; \quad \hat{\theta} = 0.22 \text{ (males)} .$$

Higher accuracy demands numerical techniques, but their use here is unnecessary.

Bearing in mind the simplicity of our model, the female fit is qualitatively quite good, even though the expected distribution is too shallow along the sides. The male data has a very pronounced local maximum at the centre, which cannot be explained by our model, and the fit is generally poor. Far better agreement between the observed and expected values for the female distribution may well be obtained if

(i) the migration rate from side colonies to their nearest interior neighbours is less than the migration rate in the reverse direction (i.e. beetles have a straightforward preference for side colonies),

(ii) the migration rate between side colonies is greater than that between interior colonies (on the flour-surface of

a well-populated vial, for example, the mass of beetles is in a highly agitated motion; see Neyman et al. p.50). These suggestions have yet to be investigated.

Within certain limits, fecundity is affected by the incidence of mating (Park (1933)). Thus fecundity will change as the total population size increases, in contrast to the fixed values taken by the parameters in our model. Moreover, it is clear from the comments of Neyman et al. that even after four months the distribution of beetles in the cube is far from being stationary, and this is contrary to the requirements of solution (5.5).

A further criticism is that we are fitting a two-dimensional model to a three-dimensional distribution. For beetles in interior cells may in fact migrate to one of six nearest-neighbours, and equations (5.2) should really be written as

$$m_{ijk}(t+1) = \omega m_{ijk}(t) = \begin{cases} (1+\lambda-\mu-6\nu)m_{ijk}(t)+\nu L_{ijk}(t) & : \text{INTERIOR} \\ (1+\lambda'-\mu'-5\nu)m_{ijk}(t)+\nu L_{ijk}(t) & : \text{SIDES} \\ (1+\lambda''-\mu''-4\nu)m_{ijk}(t)+\nu L_{ijk}(t) & : \text{EDGES} \\ (1+\lambda'''-\mu'''-3\nu)m_{ijk}(t)+\nu L_{ijk}(t) & : \text{CORNERS} . \end{cases} \tag{5.12}$$

Here $m_{ijk}(t)$ $(i,j,k = 0,\ldots,n\ ;\ t=0,1,2,\ldots)$ represents the number of beetles in colony (i,j,k) at the discrete time points t. The parameters λ', μ' ; λ'', μ'' ; λ''', μ''' now denote the birth and death rates corresponding to side, edge and corner colonies, respectively, and the function

$$L_{ijk}(t) = m_{i-1,j,k}(t) + m_{i+1,j,k}(t) + m_{i,j-1,k}(t) + m_{i,j+1,k}(t) + m_{i,j,k-1}(t) + m_{i,j,k+1}(t) \qquad (0 \leq i,j,k \leq n)$$

where we define $m_{ijk}(t) \equiv 0$ if (i,j,k) is not a lattice point (i.e. if either $i,j,k < 0$ or $i,j,k > n$) .

In view of the nature of solutions (4.10) and (5.5) for the one- and two-dimensional models, respectively, let us consider the trial solution $(i,j,k = 0,\ldots,n)$

$$(5.13) \qquad m_{ijk} = d.\cosh[(n-2i)\theta]\cosh[(n-2j)\theta]\cosh[(n-2k)\theta]$$

where d is a suitably chosen normalizing constant. Substituting (5.13) into the 'interior' equation (5.12) and simplifying the right hand side we get

$$(\omega-1-\lambda+\mu+6\nu)\cosh[(n-2i)\theta]\cosh[(n-2j)\theta]\cosh[(n-2k)\theta]$$
$$= 6\nu\cosh(2\theta)\cosh[(n-2i)\theta]\cosh[(n-2j)\theta]\cosh[(n-2k)\theta] \ .$$

Hence for (5.13) to be a solution we require that

$$(5.14) \qquad (\omega-1-\lambda+\mu+6\nu) = 6\nu\cosh(2\theta) \ .$$

If we substitute for m_{ijk} from (5.13) into the side, edge and corner equations (5.12) we obtain

$$(5.15) \quad (\omega-1-\lambda'+\mu'+5\nu) = \nu\{4\cosh(2\theta) + \cosh[(n-2)\theta]/\cosh(n\theta)\}$$

$$(5.16) \quad (\omega-1-\lambda''+\mu''+4\nu) = 2\nu\{\cosh(2\theta) + \cosh[(n-2)\theta]/\cosh(n\theta)\}$$

$$(5.17) \quad (\omega-1-\lambda'''+\mu'''+3\nu) = 3\nu\cosh[(n-2)\theta]/\cosh(n\theta) \ ,$$

respectively. These three equations are independent of the location parameters i,j and k as required. Expressions (5.14)-(5.17) all express ω as a function of θ , and so once ν and two of the four parameter differences $(\lambda-\mu)$, $(\lambda'-\mu')$, $(\lambda''-\mu'')$ and $(\lambda'''-\mu''')$ are known the values of the remaining two parameter differences are completely specified.

Now within any one layer of the Tribolium k remains

constant. In particular, when k=1 (corresponding to layer II)

(5.18) $m_{ijl} = [d.\cosh(7\theta)]\cosh[(9-2i)\theta]\cosh[(9-2j)\theta] \quad (0 \leq i,j \leq 9)$,

and if we write

$$c = d.\cosh(7\theta)$$

expression (5.18) corresponds exactly with the two-dimensional solution (5.5). Thus the data for layer II may be analyzed as though it were obtained from a two-dimensional experiment.

APPENDIX A

COVARIANCE EXPRESSION FOR THE TWO-COLONY PROCESS

Denote

$$d_1 = \alpha_1 - a_1\xi_2 + \nu_2 a_2 \quad ; \quad d_2 = \alpha_2\nu_2 - \alpha_1\xi_2$$

$$f_1 = \alpha_2 - a_2\xi_1 + \nu_1 a_1 \quad ; \quad f_2 = \alpha_1\nu_1 - \alpha_2\xi_1 \quad .$$

Then the coefficients $r_i (i=1,\ldots,6)$ in the expression

$$V_{12}(t) = r_1 e^{2\omega_1 t} + r_2 e^{(\omega_1+\omega_2)t} + r_3 e^{2\omega_2 t} + r_4 e^{\omega_1 t} + r_5 e^{\omega_2 t} + r_6$$

(2.4.15)

are given by

$$r_1 = [\lambda_1\nu_1(4a_1\omega_1^2+2d_1\omega_1+d_2)(\omega_1-\xi_2) + \lambda_2\nu_2(4a_2\omega_1^2+2f_1\omega_1+f_2)(\omega_1-\xi_1)]/\omega_1^2(\omega_1-\omega_2)^2(2\omega_1-\omega_2)$$
$$- [\nu_1 a_1(\omega_1-\xi_2) + \nu_2 a_2(\omega_1-\xi_1)]/(\omega_1-\omega_2)^2$$

$$r_2 = -2(\xi_1-\xi_2)[(a_1(\xi_1+\xi_2)^2+d_1(\xi_1+\xi_2)+d_2) - (a_2(\xi_1+\xi_2)^2+f_1(\xi_1+\xi_2)+f_2)]/\omega_1\omega_2(\omega_1+\omega_2)(\omega_1-\omega_2)^2$$
$$+ (\xi_1-\xi_2)(\nu_1 a_1-\nu_2 a_2)/(\omega_1-\omega_2)^2$$

$$r_3 = [\lambda_1\nu_1(4a_1\omega_2^2+2d_1\omega_2+d_2)(\omega_2-\xi_2) + \lambda_2\nu_2(4a_2\omega_2^2+2f_1\omega_2+f_2)(\omega_2-\xi_1)]/\omega_2^2(\omega_1-\omega_2)^2(2\omega_2-\omega_1)$$
$$- [\nu_1 a_1(\omega_2-\xi_2)+\nu_2 a_2(\omega_2-\xi_1)]/(\omega_1-\omega_2)^2$$

$$r_4 = 2[\lambda_1\nu_1(a_1\omega_1^2+d_1\omega_1+d_2)(\omega_1-2\xi_2) + \lambda_2\nu_2(a_2\omega_1^2+f_1\omega_1+f_2)(\omega_1-2\xi_1)]/\omega_1^2\omega_2(\omega_1-\omega_2)(\omega_1-2\omega_2)$$

$$r_5 = 2[\lambda_1\nu_1(a_1\omega_2^2+d_1\omega_2+d_2)(\omega_2-2\xi_2) + \lambda_2\nu_2(a_2\omega_2^2+f_1\omega_2+f_2)(\omega_2-2\xi_1)]/\omega_2^2\omega_1(\omega_2-\omega_1)(\omega_2-2\omega_1)$$

$$r_6 = [\lambda_1\nu_1 d_2\xi_2 + \lambda_2\nu_2 f_2\xi_1]/\omega_1^2\omega_2^2(\omega_1+\omega_2)$$

where

(2.3.9) $$\omega_1,\ \omega_2 = \tfrac{1}{2}[(\xi_1+\xi_2) \pm \{(\xi_1-\xi_2)^2+4\nu_1\nu_2\}^{\frac{1}{2}}] \quad .$$

APPENDIX B

SIMULATION PROGRAM FOR A ONE-WAY MIGRATION STEPPING-STONE MODEL

The program listed below is written in Fortran IV and refers to the stochastic simulation of the one-way migration model without death which is described in section 9.8.

Notation :

B - birth rate (λ)

M - migration rate (ν)

N(I) - population size of colony I-25 (I = 1,...,1000)

POP - total population size

POS - position of the wavefront

TIME - time

RAND - random integer

YFL - uniformly distributed random number on [0,1]

MAXCNT - total number of events

Program :

```
01          INTEGER COUNT,RAND,POP,POS,TAL,P
02          REAL M
03          DIMENSION N(1000),MM(25)
      C        READ IN PARAMETERS
04          READ (5,100) MAXCNT,RAND,B,M
05    100   FORMAT (2I6,2F5.1)
      C        PRINT PARAMETERS
06          WRITE (6,150) B
```

```
07    150   FORMAT (14H1BIRTH RATE = ,F9.1)
08          WRITE (6,151) M
09    151   FORMAT (18HOMIGRATION RATE = ,F5.1)
10          WRITE (6,152) RAND
11    152   FORMAT (17HORANDOM NUMBER = ,I12////)
      C           INITIAL GENERATOR RUN
12          DO 200 K=1,100
13    200   RAND = RAND*65539
      C           SET PARAMETERS
14          COUNT = 0
15          TIME = 0.0
16          TAL = 0
17          DO 210 K1=1,1000
18    210   N(K1) = 0
19          N(25) = 1
20          DO 220 K2=1,25
21    220   MM(K2) = N(K2)
22          POS = 25
23          WRITE (6,310) TAL,TIME,MM
24    310   FORMAT (I4,F8.2,25I4)
25          POP = 1
26          RATIO = B/(B+M)
      C           SIMULATE
27          DO 400 COUNT=1 ,MAXCNT
28          RATE = POP*(B+M)
29          RAND = RAND*65539
```

```
30          YFL = RAND
31          YFL = YFL*.2328307E-9+.5
32          REXP = -(ALOG(YFL))/RATE
33          TIME = TIME+REXP
      C        TYPE OF EVENT
34          RAND = RAND*65539
35          YFL = RAND
36          YFL = YFL*.2328307E-9+.5
37          TEST = YFL*POP
38          I = 25
39          X = N(25)
40    460   IF(TEST.LE.X) GO TO 500
41          I = I+1
42          X = X+N(I)
43          GO TO 460
44    500   RAND = RAND*65539
45          YFL = RAND
46          YFL = YFL*.2328307E-9+.5
47          IF(YFL.LE.RATIO) GO TO 600
48          N(I) = N(I)-1
49          N(I+1) = N(I+1)+1
50          GO TO 610
51    600   N(I) = N(I)+1
52          POP = POP+1
53          GO TO 400
54    610   IF(I.LT.POS) GO TO 400
```

```
55          POS = POS+1
56          TAL = TAL+1
57          LL = 1
58          P = POS-24
59          DO 660 L=P,POS
60          MM(LL) = N(L)
61    660   LL = LL+1
62          WRITE (6,700) TAL,TIME,MM
63    700   FORMAT (I4,F8.2,25I4)
64    400   CONTINUE
65          STOP
66          END
```

APPENDIX C

DATA ON THE SPATIAL DISTRIBUTION OF TRIBOLIUM CONFUSUM

The following data refer to the experiment of Neyman, Park and Scott (1956) and show the distribution of adult males and females within the 100 constituent small cubes in the second layer from the top of the experimental cube. Details of this experiment are given in section 11.2.

upper figure in each cell - density of females

lower figure in each cell - density of males

18.8	22.7	16.1	12.4	12.8	13.8	11.0	14.4	11.8	16.7
28.7	24.2	14.2	12.0	11.1	15.3	14.1	18.1	20.2	23.0
10.1	11.2	9.6	7.9	7.3	7.3	5.1	4.3	6.9	11.9
14.8	6.9	3.8	3.1	5.3	3.5	3.2	4.9	8.0	16.0
7.9	6.5	5.7	4.4	3.5	3.6	4.4	5.9	8.4	10.4
13.3	4.6	2.0	1.7	1.0	1.9	1.9	1.9	3.5	9.8
9.9	6.9	5.4	5.8	3.2	2.9	3.1	5.2	7.0	8.7
9.5	4.5	1.5	0.7	1.3	1.3	0.2	1.0	1.7	7.7
10.9	6.6	3.0	2.6	4.0	4.2	3.9	3.5	5.8	17.8
12.5	4.6	4.6	2.2	2.0	2.5	3.9	1.6	2.2	12.2
6.0	7.8	7.2	2.4	1.0	3.5	4.0	4.0	6.2	16.5
8.0	4.5	2.4	1.8	1.3	3.3	1.8	2.4	3.5	19.4
10.5	6.5	1.2	5.7	4.7	4.2	5.4	3.7	9.7	14.2
10.6	3.5	5.1	2.0	0.4	1.9	2.3	0.9	3.0	12.2
14.9	10.8	4.9	3.9	4.0	3.5	5.8	4.3	7.3	13.3
15.8	5.9	2.8	2.7	0.8	1.4	2.8	1.7	3.8	11.8
18.3	10.4	8.5	5.3	5.6	4.8	6.4	8.3	8.2	12.1
26.6	8.9	3.4	4.6	2.7	3.5	2.5	2.4	4.9	18.6
19.1	12.6	15.8	14.7	12.8	11.4	12.9	12.3	11.5	17.1
35.2	14.1	17.0	17.2	21.8	13.9	14.9	12.2	14.1	34.5

BIBLIOGRAPHY

1. Abramowitz, M. and Stegun, I.A. (1965). Handbook of Mathematical Functions (Dover edition). AMS 55, U.S. Department of Commerce.

2. Adke, S.R. (1964a). The generalized birth and death process and Gaussian diffusion. J. Math. Anal. Appl. 9, 336-340.

3. Adke, S.R. (1964b). A stochastic population diffusing on a finite interval. J. Ind. Statist. Ass. 2, 32-40.

4. Adke, S.R. (1969). A birth, death and migration process. J. Appl. Prob. 6, 687-691.

5. Adke, S.R. and Moyal, J.E. (1963). A birth, death and diffusion process. J. Math. Anal. Appl. 7, 209-224.

6. Alksland, M. (1975). A birth, death and migration process with immigration. Adv. Appl. Prob. 7, 44-60.

7. Andres. A. (1931). Catalogue of the Egyptian Tenebrionidae. Bull. de la Société Roy. Entom. d'Egypte 15, 74-125.

8. Arley, N and Borchsenius, V. (1945). On the theory of infinite systems of differential equations and their application to the theory of stochastic processes and the perturbation theory of quantum mechanics. Acta Math. 76, 261-322.

9. Armitage, P. (1952). The statistical theory of bacterial populations subject to mutation. J. Roy. Statist. Soc. B 14, 1-33.

10. Arnason, A.N. (1971). Migration models for animal populations. Ph.D. Thesis, University of Edinburgh.

11. Bailey, N.T.J. (1957). The Mathematical Theory of Epidemics. Griffin, London.

12. Bailey, N.T.J. (1964). The Elements of Stochastic Processes. Wiley, New York.

13. Bailey, N.T.J. (1967). The simulation of stochastic epidemics in two dimensions. Proc. Fifth Berkeley Symp. Math. Statist. Prob. Vol. 4, 237-257.

14. Bailey, N.T.J. (1968). Stochastic birth, death and migration processes for spatially distributed populations. Biometrika 55, 189-198.

15. Bartlett, M.S. (1949). Some evolutionary stochastic processes. J. Roy. Statist. Soc. B 11, 211-229.

16. Bartlett, M.S. (1960). Stochastic Population Models in Ecology and Epidemiology. Methuen, London.

17. Bartlett, M.S. (1966). An Introduction to Stochastic Processes (second edition). Cambridge Univ. Press.

18. Bellman, R. (1953). Stability Theory of Differential Equations. McGraw-Hill, New York.

19. Bharucha-Reid, A.T. (1960). Elements of the Theory of Markov Processes and Their Applications. McGraw-Hill, New York.

20. Billard, L. (1973). Factorial moments and probabilities for the general stochastic epidemic. J. Appl. Prob. 10, 277-288.

21. Broadbent, S.R. and Kendall, D.G. (1953). The random walk of Trichostrongylus retortaeformis. Biometrics 9, 460-466.

22. Canosa, J. (1973). On a non-linear diffusion equation describing population growth. IBM J. Res. Develop. 17, 307-313.

23. Chamberlain, A.C. (1970). Deposition and uptake of windborne particles. Nature 225, 99-100.

24. Chiang, C.L. (1968). Introduction to Stochastic Processes in Biostatistics. Wiley, New York.

25. Cox, D.R. and Miller, H.D. (1965). The Theory of Stochastic Processes. Methuen, London.

26. Cox, D.R. and Smith, W.L. (1957). On the distribution of Tribolium Confusum in a container. Biometrika 44, 328-335.

27. Crow, J.F. and Kimura, M. (1970). An Introduction to Population Genetics Theory. Harper and Row, New York.

28. Crump, K.S. (1970). Migratory populations in branching processes. J. Appl. Prob. 7, 565-572.

29. Daniels, H.E. (1952). The statistical theory of stiff chains. Proc. Roy. Soc. Edinb. 63, 290-311.

30. Davies, W.K., Lewis, G.B. and Randall, H.A. (1968). Some distributional features of the Foot and Mouth epidemic. Nature 219, 121-125.

31. Davis, A.W. (1965). On the theory of birth, death and diffusion processes. J. Appl. Prob. 2, 293-322.

32. Davis, A.W. (1967a). Branching-diffusion processes with no absorbing boundaries. I. J. Math. Anal. Appl. 18, 276-296.

33. Davis, A.W. (1967b). Branching-diffusion processes with no absorbing boundaries. II. J. Math. Anal. Appl. 19, 1-25.

34. Davis, A.W. (1970). Some generalizations of Bailey's birth, death and migration model. Adv. Appl. Prob. 2, 83-109.

35. Deevey, E.S. (1947). Life tables for natural populations of animals. Quart. Rev. Biol. 22, 283-314.

36. Dietz, K. (1967). Epidemics and rumours : a survey. J. Roy. Statist. Soc. A 130, 505-528.

37. Downton, F. (1968). The ultimate size of carrier-borne epidemics. Biometrika 55, 277-289.

38. Feller, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. Acta Biotheoretica 5, 11-40.

39. Feller, W. (1966). An Introduction to Probability Theory and its Applications. Vol. 2. Wiley, New York.

40. Fisher, R.A. (1936). The wave of advance of advantageous genes. Ann. Eugen. 7, 355-369.

41. Gani, J. (1965). Stochastic models for bacteriophage. J. Appl. Prob. 2, 225-268.

42. Gani, J. and Yeo, G.F. (1965). Some birth-death and mutation models for phage reproduction. J. Appl. Prob. 2, 150-161.

43. Griffiths, D.A. (1972). A bivariate birth-death process which approximates to the spread of a disease involving a vector. J. Appl. Prob. 9, 65-75.

44. Halliday (1929). (in the discussion on Mr Soper's paper). J. Roy. Statist. Soc. A 92, 67-69.

45. Harris, T.E. (1963). The Theory of Branching Processes. Springer-Verlag, Berlin.

46. Helland, I.S. (1975). The condition for extinction with probability one in a birth, death and migration process. Adv. Appl. Prob. 7, 61-65.

47. Ince, E.L. (1956). Integration of Ordinary Differential Equations (seventh edition). Oliver and Boyd, Edinburgh.

48. Kamke, E. (1959). Differentialgleichungen; Lösungsmethoden und Lösungen, I. Akadem. Verl.-G., Leipzig.

49. Karlin, S. and McGregor, J.L. (1957a). The differential equations of birth and death processes and the Stieltjes moment problem. Trans. Amer. Math. Soc. 85, 489-546.

50. Karlin, S. and McGregor, J.L. (1957b). The classification of birth and death processes. Trans. Amer. Math. Soc. 86, 366-400.

51. Kendall, D.G. (1948a). On the generalized 'birth-and-death' process. Ann. Math. Statist. 19, 1-15.

52. Kendall, D.G. (1948b). On the role of variable generation time in the development of a stochastic birth process. Biometrika 35, 316-330.

53. Kendall, D.G. (1950). An artificial realization of a simple 'birth-and-death' process. J. Roy. Statist. Soc. B 12, 116-119.

54. Kendall, D.G. (1965). Mathematical models of the spread of infection : Mathematics and Computer Science in Biology and Medicine. London, H.M. Stationery Office, 213-225.

55. Kimura, M. (1953). 'Stepping stone' model of population. Ann. Rept. Nat. Inst. Genetics, Japan 3, 62-63.

56. Kimura, M. and Weiss, G. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49, 561-576.

57. Kingman, J.F.C. (1969). Markov population processes. J. Appl. Prob. 6, 1-18.

58. Kolmogorov, A.N., Petrovsky, I.G. and Piscounov, N.S. (1937). Étude de l'equation de la diffusion avec croissance de la quantité de matière et son application à une problème biologique. Bull. de l'Univ. d'État à Moscou (sér. intern.). Sect. A 1, fasc. 6, 1-25.

59. Marris, R.L. (1970). The Economic Theory of 'Managerial' Capitalism. Macmillan, London.

60. Maruyama, T. (1969). Genetic correlation in the stepping stone model with non-symmetrical migration rates. J. Appl. Prob. 6, 463-477.

61. Maruyama, T. (1970). Stepping stone models of finite length. Adv. Appl. Prob. 2, 229-258.

62. Matis, J.H. and Hartley, H.O. (1971). Stochastic compartmental analysis : model and least squares estimation from time series data. Biometrics 27, 77-102.

63. May, R.M. (1973). Stability and Complexity in Model Ecosystems. Princeton Univ. Press.

64. Maynard Smith, J. (1974). Models in Ecology. Cambridge Univ. Press.

65. Mode, C.J. (1971). Multitype Branching Processes. American Elsevier, New York.

66. Mollison, D. (1972a). The rate of spatial propagation of simple epidemics. Proc. Sixth Berkeley Symp. Math. Statist. Prob. Vol. 3, 579-614.

67. Mollison, D. (1972b). Possible velocities for a simple epidemic. Adv. Appl. Prob. 4, 233-257.

68. Montroll, E.W. (1967). On non-linear processes involving population growth and diffusion. J. Appl. Prob. 4, 281-290.

69. Neyman, J., Park. T. and Scott, E.L. (1956). Struggle for existence. The Tribolium model : biological and statistical aspects. Proc. Third Berkeley Symp. Math. Statist. Prob. Vol. 4, 41-79.

70. Norris, K.P. and Harper, G.J. (1970). Windborne dispersal of Foot and Mouth virus. Nature 225, 98-99.

71. Park, T. (1933). Studies in population physiology. II. Factors regulating initial growth of Tribolium Confusum populations. Jour. Exp. Zool., 65, 17-42.

72. Piaggio, H.T.H. (1962). An Elementary Treatise on Differential Equations and Their Applications. Bell, London.

73. Pielou, E.C. (1969). An Introduction to Mathematical Ecology. Wiley, New York.

74. Puri, P.S. (1968). Interconnected birth and death processes. J. Appl. Prob. 5, 334-349.

75. Raman, S. and Chiang, C.L. (1973). On a solution of the migration process and the application to a problem in epidemiology. J. Appl. Prob. 10, 718-727.

76. Ramsey, A.S. (1960). Statics. Cambridge Univ. Press.

77. Renshaw, E. (1970). Birth, death and migration processes. M. Phil. Thesis, University of Sussex.

78. Renshaw, E. (1972). Birth, death and migration processes. Biometrika 59, 49-60.

79. Renshaw, E. (1973a). Interconnected population processes. J. Appl. Prob. 10, 1-14.

80. Renshaw, E. (1973b). The effect of migration between two developing populations. Proc. 39th Session Intern. Statist. Inst. 2, 294-298.

81. Renshaw, E. (1974). Stepping stone models for population growth. J. Appl. Prob. 11, 16-31.

82. Renshaw, E. (1975). Spatial dynamics of biological populations. Adv. Appl. Prob. 7, 459-460.

83. Riordan, J. (1958). An Introduction to Combinatorial Analysis. Wiley, New York.

84. Ruben, H. (1962). Some aspects of the emigration-immigration process. Ann. Math. Statist. 33, 119-129.

85. Severo, N.C. (1967). Two theorems on solutions of differential-difference equations and applications to epidemic theory. J. Appl. Prob. 4, 271-280.

86. Severo, N.C. (1969a). A recursion theorem on solving differential-difference equations and applications to some stochastic processes. J. Appl. Prob. 6, 673-681.

87. Severo, N.C. (1969b). The probabilities of some epidemic models. Biometrika 56, 197-201.

88. Sherman, B. (1956). The limiting distribution of Brownian motion on a finite interval with instantaneous return. Westinghouse Research Laboratory Scientific Paper 60 - 94698 - 3 - P3.

89. Skellam, J.G. (1951). Random dispersal in theoretical populations. Biometrika 38, 196-218.

90. Stakman, E.C. and Harrar, J.G. (1957). Principles of Plant Pathology. Ronald Press, New York.

91. Tinline, R. (1970). Lee wave hypothesis for the initial pattern of spread during the 1967-68 Foot and Mouth epizootic. Nature 227, 860-862.

92. Usher, M.B. and Williamson, M.H. (1970). A deterministic matrix model for handling the birth, death and migration processes of spatially distributed populations. Biometrics 23, 1-12.

93. Waugh, W.A.O'N. (1972). Taboo extinction, sojourn times, and asymptotic growth for the Markovian birth and death process. J. Appl. Prob. 9, 486-506.

94. Weiss, G. and Kimura, M. (1965). A mathematical analysis of the stepping stone model of genetic correlation. J. Appl. Prob. 2, 129-149.

95. Wiggins, A.D. (1960). On a multicompartment migration model with chronic feeding. Biometrics 16, 642-658.

96. Zeldovitch, Y.B. (1951). Theory of Flame Propagation. Washington D.C., NACA T.M. 1282.

97. Ziegler, P. (1969). The Black Death. Collins, London.